Evaluating aid Effectiveness in the aggregate
a critical assessment of the evidence

Dalgaard, Carl-Johan Lars; Hansen, Henrik

Publication date:
2010

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Dalgaard, C-J. L., & Hansen, H. (2010). *Evaluating aid Effectiveness in the aggregate: a critical assessment of the evidence*. (1 ed.) Ministry of Foreign Affairs of Denmark. Danida. Evaluation studies Vol. 2010 No. 1

Ministry of Foreign Affairs of Denmark
Danida

FROM DANIDA'S EVALUATION DEPARTMENT – DEVELOPMENT COOPERATION, FEBRUARY 2010

EVALUATING AID EFFECTIVENESS IN THE AGGREGATE: A CRITICAL ASSESSMENT OF THE EVIDENCE

EVALUATION STUDY 2010/1

Evaluating Aid Effectiveness in the Aggregate: A critical assessment of the evidence

January 2010

Carl-Johan Dalgaard, Department of Economics, University of Copenhagen
Henrik Hansen, Institute of Food and Resource Economics, University of Copenhagen

---

*Acknowledgement: We are grateful to Arvind Subramanian for sharing the data from Rajan and Subramanian (2008). Disclaimer: The views expressed are those of the authors and do not necessarily represent the views of the Ministry of Foreign Affairs of Denmark. Errors and omissions are the responsibility of the authors.

Contents

Executive Summary

The purpose of the present evaluation study is to discuss the empirical studies that attempt to estimate the impact of foreign aid on economic growth. The study draws on a previous evaluation study (Dalgaard and Hansen, 2009), which introduces the general econometric methodology involved in making assessments about the aggregate impact of aid. In order to fully benefit from the discussion below it is therefore advisable for readers without prior knowledge of econometrics to review the material discussed in Dalgaard and Hansen (2009).

The present study provides insights into the following questions

1. What are the central mechanisms linking aid to growth?
2. How much should one expect from aid *a priori*?
3. What are the best available estimates of the impact from total aid on economic growth in income per capita?
4. Does aid modality matter?

Key points:

- On *a priori* grounds a 1% increase in the recipient aid-to-gdp ratio for equipment investment should – at best- raise the long-run level of GDP per capita by 0.5-0.67%
- On *a priori* grounds a 1% increase in the recipient aid-to-gdp ratio for infrastructure investments should – at best – increase the level of GDP per capita by 0.23-0.45%
- Investments in human resources (health, schooling and fertility) also holds clear potential to increase GDP per capita. But the effects will only materialize very slowly over time
- For total aid: Best available empirical estimates suggest an elasticity of up to 0.3%; zero cannot be rejected
- For aid modalities: There is some evidence that project aid is relatively more effective that other modalities, but the evidence is inconclusive
- Overall, there is little consensus on the impact from aid on growth. The problem of identifying the *causal* impact of aid on growth remains largely unsolved.

1. Introduction

This evaluation study complements an earlier study of the methodological issues involved in examining the aggregate impact from aid on growth (Dalgaard and Hansen, 2009). In that study we particularly focused on why aid effectiveness almost always is judged from aid's impact on growth in income per capita, why standard statistical regression analysis (so-called ordinary least squares regressions) often are misleading in terms of providing information of the *causal* effect of aid on growth, and finally, what one might do in order to identify the impact of aid (so-called two stage least squares regressions). In the present study we take this material as given, and proceed to evaluate the empirical literature on the topic.

We focus on two strands of literature. The first line of literature, which is the most fully developed, focuses on assessing the impact from *total* aid disbursements. This literature is by now vast. As a consequence a full survey of all contributions in the area will not be attempted here. We have therefore chosen to carefully discuss the most influential contributions to the literature from the 21st century (based on citations), and in addition contributions that have recently appeared in leading academic journals. Common features of this list of studies is that: (a) they examine the impact of aid on *growth* in income per capita, and (b) they all (at least *try* to) deal with the methodological difficulties outlined in Dalgaard and Hansen (2009). A more recent strand of literature examines aid modalities. This literature is far smaller, which allows us to survey it more fully. As a consequence we attempt to survey it more fully.

The present study is structured as follows. In the first two sections we discuss why aid arguably should impact on growth, and how big an effect from aid one *a priori* should expect. This provides a clear yardstick against which we can evaluate empirical estimates. We survey the empirical literature in Section 3. As noted we focus on the most recent and influential contributions. More specifically, we discuss contributions which strive to identify the causal impact of total aid on growth, as well as some recent work which attempts to gauge the impact of aid modalities. A final section contains concluding remarks.

2. How Aid Affects Growth

As a point of departure it is useful to clarify how GDP is determined at the macro level, seen through the lens of economic growth theory. Figure 1 provides an overview over the main *proximate* determinants of GDP. Three broad determinants of production (GDP) can be distinguished: Physical Capital, Human Resources and "Productivity".

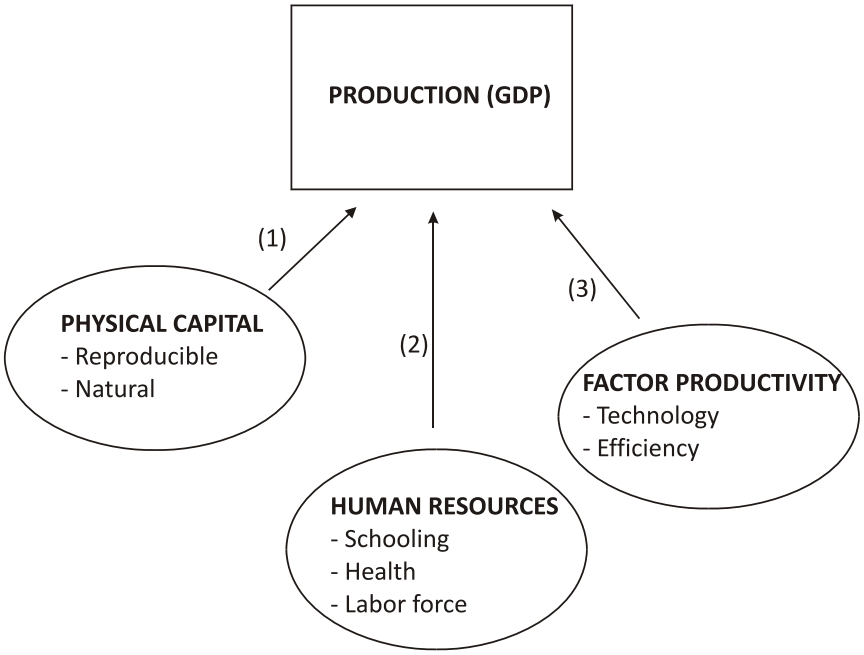

Figure 1. Accounting for GDP: Proximate Sources of Growth.

The first determinant, "physical capital", includes equipment (i.e., "reproducible capital") as well as land ("natural capital"). "Human resources", in turn, is determined by the size of the labor force as well as how educated and healthy the labor force is. Finally, the level of GDP is determined by "productivity", which depends on technology and aggregate "efficiency". The latter rises if scarce resources, to an increasing extent, are allocated towards the most productive sectors. It also depends on the supply of public goods like infrastructure. If foreign aid is to have an impact on economic growth it must ultimately be able to stimulate the rate at which these three categories of "inputs" expand.

This way of characterizing the growth process may seem overly simplistic. Where are the demand considerations? In this respect it is important to appreciate that economists typically

subscribe to the view that "long-run" growth is driven mainly from the *supply side*.[1] That is, growth is determined by the rate of expansion of the inputs highlighted in Figure 1. In the shorter run, however, output may drift away from its "potential", as determined by the supply side. This is where demand stimulus may affect GDP growth. But once the economy has reached the potential level of output, further stimuli will not be able to increase growth; it will only result in inflation. In the present context we are preoccupied with growth in the long-run, for which reason we will be focusing on the ability of foreign aid to stimulate input growth, or "potential" output.[2]

Another point worth elaborating is the statement that Figure 1 depicts *proximate* sources of growth. This means that in *accounting* for growth in GDP we can attribute past growth to the highlighted factors: capital, human resources, and productivity. But that is not the same as *explaining* growth. In order to explain growth we need to understand *why* individual inputs are expanding at the rate they are.

A final point worth making is that foreign aid may affect growth via the same channel in multiple ways. For a proper assessment of the scope for aid in spurring (or hindering) growth all these pathways needs to be taken into account. We need to understand, to the greatest extent possible, the impact of aid once the reaction in all areas (markets) of the economy has been taken into account. In the terminology of economics: we need to understand the "general equilibrium effects" of aid. A few examples may be useful: Capital input expands as a result of domestic investments. Basically there a two ways domestic investment may increase; either through domestic resource mobilization (domestic savings) or via foreign contributions (foreign direct investments (FDI) and/or foreign aid). Hence, foreign aid could impact on this growth determinant directly. Aid funded investments in the production sector is an example of how aid would stimulate capital accumulation. However, even if such investments are successful they may not increase the over-all capital stock if capital is internationally mobile. If foreign aid and FDI are *substitutes*, it is possible that aid (in the form of capital transfers) simply crowds out FDI, leaving the net effect on domestic capital accumulation ambiguous. If aid is *complementary* to FDI, however, the opposite is the case.

---

[1] What's "the long-run"? In the present context economists usually have a time horizon of at least a decade in mind.

[2] That is not to say that demand is completely irrelevant from a long-run perspective. Importantly, it may influence the incentive to innovate and thus the speed of technological progress. Hence, demand will influence "input growth" *indirectly*; a complete dichotomy between the demand side and the supply side is of course not viable.

As another example, consider initiatives that improve health. For instance, aid financed projects that help distribute medicine, or health care services more broadly. The direct effect on "health" from such an initiative is likely positive; taken in isolation, aid induced health improvements should stimulate labor productivity (cf Figure 1). However, better health also stimulates survival. From a welfare perspective this is a good thing, of course. But, as more people survive population density goes up, which could have adverse economic effects. Greater population density may well *lower* average labor productivity in subsistence societies, as more people are forced to subside on the same plot of land. If average productivity declines, so will income and thus nutrition. The latter could have an off-setting indirect effect on the over-all health status of the population.

Hence, in order to form priors about aid effectiveness a suitable framework capturing the general equilibrium effects of aid inflows need to be invoked. In the next section we examine the quantitative significance of each of the channels from Figure 1. It is worth bearing in mind that the empirical literature on aid effectiveness in fact purports to estimate the over-all, or "general equilibrium", impact of aid. Without a prior about what to expect from aid, the parameter estimates from such studies are hard to come to grips with. This is why we begin our review of the econometric literature on aid effectiveness with an excursion into theory-based calculations (or, in the terminology of the field: *calibration*) of aid effectiveness. The goal is to provide a sense of magnitudes: What is the plausible impact from aid on growth?

2.1. Channel 1: Physical Capital

The most used framework for analyzing the growth process remains the celebrated Solow (1956) model. While later contributions added rigor, the basic model remains at the core of useful macroeconomic models, and continues to be used in academic research as a central organizing framework. An important reason why the model is so popular is that it has proven to hold considerable explanatory power in understanding growth and income differences across the world (e.g., Mankiw et al., 1992; Caselli et al., 1996; Hoeffler, 2002).

The Solow model distils the logic of an important dynamic multiplier process, with bearing on capital accumulation and long-run labor productivity. The process works as follows.

A small increase in income per capita increases savings per capita as well as investments per capita. The former holds under the assumption that households save a constant (possibly arbitrarily small) fraction of their income; if households' access to formal credit markets is limited this is not a bad assumption. The equality between savings and investments hold by

identity *in a closed economy* (more on this below).[3] Now, higher investments per capita instigates capital accumulation; the acquisition of more equipment per employee, by firms. This in turn implies that per capita production and income expands, which facilitates further savings, investment and capital accumulation. In the end, however, the process is bounded. The reason is found in the *diminishing returns* principle: an increase in capital (per unit of labor) raises income, but progressively by less the more capital is already in place. In the long run, therefore, income per capita settles down at a constant level.[4]

Gauging the "power" of channel one in Figure 1 is equivalent to asking how powerful the above multiplier process is. In the end, the size of the multiplier depends crucially on how quickly diminishing returns set in. That is, how strongly marginal productivity of capital declines as the stock of the same goes up. This tendency can be pinned down quantitatively, after which the multiplier can be assessed.

The result can be boiled down to a simple rule of thumb: if the savings rate increases by 1 percent, this will – in the long-run -- induce an increase in labor productivity (thus, loosely speaking GDP per capita) of between 0.5 percent and 0.67 percent. The size of this multiplier should be sobering; it is not large.

To illustrate we can calculate, roughly, what we should have expected from 30 years of aid efforts on the Sub-Sahara African continent, *had all aid been spend on capital accumulation*.[5]

---

[3] As a matter of national accounts identity (i.e., something which *always* holds), GDP may be spend in two ways in a closed economy: Either it is used for investment, or for consumption (private, or public). Equivalently, the accounting identity tells us, that GDP net of consumption must equal total investment. But GDP net of consumption is precisely total savings.

[4]Unless technology progresses. Technological progress works so as to off-set diminishing returns, by making machines more productive. As a result, technological change can ensure the economy keeps growing, as has been the case in most rich countries during the last (roughly) 150 years.

[5] The following discussion draws on Dalgaard and Erickson (2009).

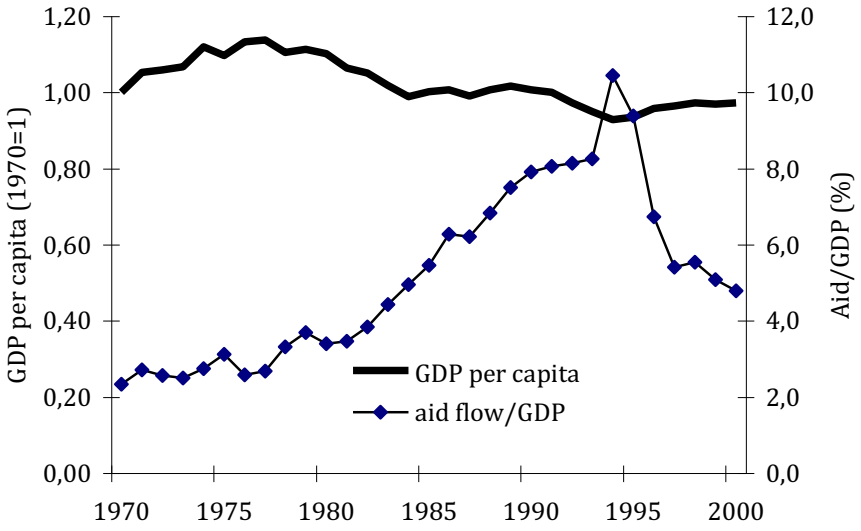

Figure 2. Evolution of Aid/GDP and GDP per capita: Sub-Sahara Africa 1970-2000.
Source: Dalgaard and Erickson (2009).

Figure 2 depicts the evolution of GDP per capita in Sub-Sahara Africa (SSA) as well as the time path of the aid to GDP ratio. In the figure GDP for the continent as well as aid flows have been pooled. In effect, therefore, we treat the entire region as "one big country". As is plain to see, growth has been broadly absent on average, whereas aid inflows have increased over the period. The average aid-to-GDP ratio for the period as a whole is about five percent, starting in 1970 at about two percent.

Using the Solow model we can now consider an experiment whereby we increase aid flows as a fraction of GDP by three percentage points; the difference between aid's initial level and the average for the 1970-2000 period. Factoring in domestic savings on the continent (on average 12 percent), the question is how much growth an increase in investment, from 14 to 17 percent will "buy"?

Calculations show that the *level* of GDP per capita should rise by between 9 and 14 percent, depending on whether the multiplier is 0.5 or 0.67. This amounts to an estimated acceleration in the average growth rate of GDP per capita over the period by between 2/10 and 4/10 of a percentage point.

While these calculations show that aid inflows (in the amounts given) should only have had a relatively minor impact on growth in SSA (had all flows been spend on physical capital) it is important to stress that it is in all likelihood *an upper boundary* to what can be expected from aid via this channel.

The size of the impact from aid flows that we have calculated so far hinges on several unrealistic assumptions. First, we have assumed that the 30 year period is long enough to "harvest" the full long-run effect from investments. Second, we have assumed that the economies can be viewed as completely closed, aside from aid. Third, *all* aid inflows are productive and were used for capital accumulation with *zero* waste. We shall return to the second assumption momentarily. For now, it is worth remarking on the first assumption. [6]

If one reexamines the multiplier process above, it is clear it implicitly requires time. Hence, the "full effect" of an increase in the savings rate is not obtained instantaneously. Rather, the process is lengthy. Best available estimates for SSA suggests that after a shock the economy closes only four percent of the "gap" to its long-run level each year absent new shocks to the economy (Hoeffler, 2002). Taking this into account brings down the level estimate above to between six and ten percent, after 30 years. Equivalently, the implied "growth acceleration" falls to between (slightly less than) 2/10 and 3/10 of a percentage point. This, however, is still an upper boundary; in part because not all aid investments are productive, and in part because of the closed-economy assumption.

Next we turn to the second potentially problematic assumption. The calculations assume that all capital accumulation in SSA is fueled either by aid or domestic savings. This is not the case. From 1970 to 2006 foreign direct investments (FDI) rose from about one percent of GDP to roughly three percent of GDP (Sunesen, 2008). As a result, it is not accurate (especially during the last few decades) to think of SSA a "closed". What are the implications of "opening" the economy to capital mobility, for the above multiplier?

The answer is that the multiplier, with respect to labor productivity, falls to zero. The reason is that, in a global economy, capital will seek out the place where the return is the highest. If the return on investments in Africa declines, as a result of aid investments and diminishing returns in combination, private investors will look elsewhere when contemplating new initiatives, or they may simply relocate. This will, over time, tend to bring down the "African return" until it is once again realigned with what can be obtained elsewhere. Evidence of this sort of arbitrage is to be found in the data (Caselli and Feyrer, 2007). In the context of "aid capital investments" the implication is that aid inflows no longer will increase the *total* per capita capital stock in a society; aid simply crowds out private flows. In a recent study Selaya

[6] As for the third assumption; "waste" is to be expected in practice. Sometimes investments simply do not work out no matter how well planed out they may be. Moreover, it is hard to rule out that part of the flows could be wasted in more insidious ways; being expropriated by corrupt rulers and bureaucrats, for instance. Whatever its source, this kind of waste would reduce the impact of aid roughly proportionally. See Dalgaard and Erickson (2009).

and Sunesen (2008) find that, in fact, aid *used for capital accumulation* has -- on average -- led to a 1:1 crowding out of private capital (FDI).

It should be clear from these considerations that while aid may fuel capital accumulation and economic growth the impact is at best a minor one. SSA is the most "aided" continent on the planet. And yet, under the best of circumstances, the total impact from these aid efforts over the period 1970-2000 should not have increased GDP per capita appreciably. At worst one should have expected close to zero impact on labor productivity in the aggregate. That is, if all flows had been used to stimulate capital accumulation. [7]

These conclusions may seem surprisingly pessimistic. In other studies the impact of aid financed capital investment is deemed to be rather substantial. A recent example is the study by Deverajan et al. (2002), which formed part of the scientific basis for calculating the "price tag" on the attainment of the first Millennium Development Goal (MDG): cutting world poverty (as of 1990) in half by 2015. The authors calculate that extreme poverty can be cut in half by 2015 by spending between US$ 40-70 billion per year on capital investment. While in absolute terms this seems like a large number, it should be borne in mind that US$ 40-70 billion can be mobilized by increasing aid donations by only a few tenth of a percentage point of total donor (DAC) GDP. This somewhat modest monetary effort should, according to the calculations of Deverajan et al. (2002), be able to pull hundreds of millions of people out of poverty.

The analysis conducted is not unusual for "policy research" in the donor community (Easterly, 1999). But it should be stressed that the underlying economic model, basically the so-called "two-gap model" (Chenery and Stout, 1966), is highly problematic for present purposes.

For starters, this model assumes that there are *no* diminishing returns to capital investment. There is now overwhelming evidence to the contrary. In addition, the model assumes that all gains in terms of growth are realized *instantaneously* upon the infusion of aid-investment. Again, this assumption is not empirically meaningful.

For these reasons the modeling approach adopted by many aid practioners overestimates the impact of aid *dramatically*, forming a poor basis for policy advice.[8] By implication it also

---

[7] Naturally, this does not contradict evidence of successful aid financed investment endeavors at the *micro* level. But micro level experiences do not always translate one-for-one into aggregate outcomes, as should now be clear.

[8] See Easterly (1999) for a thorough discussion and devastating critique.

underestimates aid requirements tremendously; for instance in the context of the first MDG.[9]

So why is the "two-gap" model still being used by aid practioners? An often heard argument is that the model is simple, and thus represents a convenient tool for "back-of-the-envelope" calculations of aid requirements. In reality, however, if one uses a Solow model as an organizing framework the computing strain is not increased significantly (Dalgaard and Erickson, 2009). It is about time the two-gap framework is finally abandoned by aid practioners.

2.2. Channel 2: Human Resources

The notion that a better educated worker is a more productive worker has a very long history in economics. Empirically, more years of schooling does seem to lead to higher wages, at the level of the individual. Similarly there is amble evidence that health status improves individual productivity. As a result there is good reason to believe that both of these dimensions of human resources can stimulate growth in per capita GDP.

The same is not true for the last component of the "human resource" category: The labor force. In societies where land is a key input factor (and usually conceived as more or less fixed) a larger labor force will work to *reduce* average productivity. In more advanced societies, where land is less important in production, growth in the labor force will have a similar, albeit smaller, effect. When labor force growth increases existing capital equipment is spread more thinly across employees, which also reduces labor productivity. The latter effect is usually referred to as the "capital diluting effect" of population growth.

A key stumbling block for a full understanding of total impact of each "human resource" variable is that they are mutually highly interconnected. In societies featuring high fertility average investments in child education and health (i.e., nutrition), tends to be lower. In turn, the perceived costs and benefits to education and nutrition will affect fertility as well. Hence, all three components are undoubtedly the joint outcome from household decisions. Initiatives that affect fertility are therefore likely to affect health and schooling also, and vice versa.

For instance, if schooling is perceived by the parent to be a (costly) necessity for a child to have a successful life, parents may react by lowering fertility so as to be able to afford such investments; the perceived benefits of an education (pecuniary or otherwise) should thus

[9]See Dalgaard and Erickson (2009) for alternative calculations using the Solow framework as underlying growth model.

influence fertility. Conversely, in high fertility regions physical capital tends to be a more of a scarce resource relative to labor. As higher levels of physical capital tend to raise the remuneration for high skilled individuals (reflecting, so-called "capital-skill complementarity"), higher fertility will work to lower the incentive to pursue an education.

Hence, in order the gauge the total impact from these individual channels on growth, a model framework needs to be invoked that ensures the interaction between the human resource variables is taken into account. A fully encompassing framework does not exist, unfortunately. Still, recent research has made progress. In the following two sub-sections we evaluate the likely effect from (i) interventions aimed at stimulating human capital, and (ii) interventions aimed at lowering fertility.

2.2.1. Human capital

Ashraf, Lester and Weil (2008) examine the impact from changes in longevity (thus stimulating the health component of human capital) by simulating a model-economy. Essentially, they augment a Solow model (as discussed above) by introducing health and schooling as productive inputs in production. They also admit health to influence schooling, and finally, they introduce a careful modeling of demographics.

The fundamental question they seek to answer is the following. Suppose an economy faces a major change in longevity; life expectancy at birth increases from 40 to 60 years. What is the impact on economic growth? This particular experiment is not arbitrary but mimics magnitudes observed in the context of the "international epidemiological transition" in the 1940s, associated with the discovery and diffusion of penicillin and DDT. For the purpose of gauging potential aid effectiveness this is a useful experiment to examine; a substantial amount of aid spending goes into health initiatives. As a result, the present experiment provides a sense of what such initiatives (at best) will be able to accomplish. In addition, the experiment involves changes in all three human resource variables simultaneously, allowing one to evaluate the "net effect" from stimulating them all. In this sense the experiment may also be indicative of what to expect if education was stimulated, which easily could involve *indirect* effects on health and population size.

Turning to the results of the analysis one might *a priori* expect that (medical) health innovations leads to faster economic growth; especially in places with a high disease burden. Inherently penicillin makes the population healthier which should stimulate labor productivity. In addition, healthier individuals tend to obtain more education, which also stimulates productivity. These effects could instigate a virtuous circle, when interacted with the "capital multiplier" discussed above. Namely, higher income (due to health and

schooling) implies greater savings, investments, capital accumulation and thus higher income still. This mechanism is indeed operative in the model of Ashraf et al. (2008). But the above mechanisms are not the end of the story.

The difficulty is that better health also instigates population growth for an extended period of time. In the simulations population growth is stimulated for two (somewhat mechanical) reasons. First, when longevity rises, the population automatically expands as deaths are avoided. Second, some of the people who, without treatment, would have died are young women. After the intervention they may live out their reproductive life, thus giving birth to more children. This too stimulates population growth.[10] In the end, the question is: what does the resulting trajectory for GDP per capital look like? The main result is that the health intervention *does* increase GDP per capita, but only in the *very* long run.

Eventually the above intervention increases GDP per capita by about 15%, according to the simulations performed by Ashraf et al. (2008); this is the joint result from expanding health, schooling and long-run population size. In this sense this experiment provides an assessment of the strength of "channel 2" in its totality.

Two things are worth observing though. First, while 15% is a very respectable effect the experiment involves a *major* intervention; an increase in longevity by 20 years. Second, the gains are obtained only in the "very" long-run, in the sense that GDP per capita only rises many decades after the intervention. During the first *four* decades after the intervention *no* gains in GDP per capita are realized. In fact, GDP per capita declines during the first several decades until the productivity gains start to emerge.[11] Hence, initially the productivity diluting effects of population growth dominates, though productivity from better health and more schooling (along with the indirect gains via the capital multiplier) ultimately will raise livings standards.

These results all relate to a closed economy setting. What happens if we allow for capital mobility? In the event the economy has access to international capital, the initial decline in GDP per capita is *not* realized. The reason is that more labor input, other things being equal,

---

[10] The analysis of Ashraf et al. (2008) does not consider household optimization. Hence, it does not take into account that changes in longevity could change fertility choices of households.

[11] A recent careful study by Acemoglu and Johnson (2007) examines the issue by way of an econometric approach to the issue at hand. The authors find a modest negative impact on GDP per capita from increases in longevity over a 60 year horizon. Hence, their results are actually somewhat more "pessimistic" than the simulation study by Ashraf et al. (2008). Recently, several papers have criticized the Acemoglu and Johnson (2007) study, and claim a positive effect of increasing longevity, at least in some countries (Bloom et al., 2009); Cervellati and Sunde, 2009).

tends to increase capital productivity which works to "draw in" foreign capital (FDI) off-setting the capital diluting effect of population growth. Another benefit from allowing capital mobility is that the long run productivity effect from the health intervention may be enlarged, since better health and schooling also leads to "crowding in". In the end the long-run impact of the intervention may be as large as 25%. Again, however, this gain is *not* obtained for a long time; it takes 60 years before the 25% gain is obtained.

It is interesting to observe, that while opening the economy to private capital flows would *reduce* the impact of aid in the context of physical capital accumulation, the opposite is true in the case of human resources. Again, the intuition is that better educated (or healthier) workers tend to make "machines more productive"; it stimulates the return on capital investment which works to attract FDI. The study by Selaya and Sunesen (2008) provides some support for this mechanism, by demonstrating that aid used e.g. on human capital investments tends to draw in FDI.[12]

The main message from the study by Ashraf et al. (2008) is that one needs to be cognizant of the fact that large gains in GDP per capita are not easily obtained by way of health initiatives. Stimulating health and schooling does increase income. But not over night, and the impacts are in all likelihood modest in comparison to international cross-country productivity differences (see Dalgaard and Hansen, 2009)).

Before we turn to fertility some remarks on schooling are called for. As alluded to above, it is well established in the literature that a more educated labor force is a more productive labor force. Again, the exact magnitudes can be hard to assess in the aggregate, since schooling may influence fertility as well as mortality. But a plausible *pure* effect of schooling is that one additional year of schooling in the labor force on average will raise the long-run level of GDP per worker by about 10%. It is once again worth bearing in mind the time required; increasing school enrolment rates in the population will not show up in higher labor productivity immediately, but requires decades to leave its mark.

2.2.2. Fertility

We next turn to the plausible impact of initiatives targeted *directly* at fertility. If indeed population growth is detrimental to economic growth these initiatives could have a substantial direct economic impact. Indirectly, they may also foster increased investments in

[12] At least insofar as aid disbursements on "complementary" (to capital) inputs is relatively modest. At high levels the opposite is true. An interpretation of this finding is that more inputs increases income per capita which instigates more domestic savings which tends to displace FDI by the same logic that aid may crowd out FDI.

schooling and health. Whereas the impact of health investments (or human capital investments) may be reduced by increased fertility, initiatives aimed at reducing fertility may therefore be regarded *a priori* as more promising.

Figure 3 shows the distribution of fertility across the world in 2007. It is plain to see that fertility varies enormously across countries. In the 25 percent of the countries with the lowest fertility rate women give birth to less than 1.8 children during their life span whereas women in the 25 percent of the countries with the highest fertility rate give birth to more than 3.8 children; a difference of two children. These differences translate into differences in population growth, as indicated by Figure 4. The figure illustrates the correlation between the fertility rate and the crude birth rate. The crude birth rate is easy to translate into population growth, since the latter is simply the difference between the numbers of children born per 100 citizens net of annual deaths per 100 citizens.

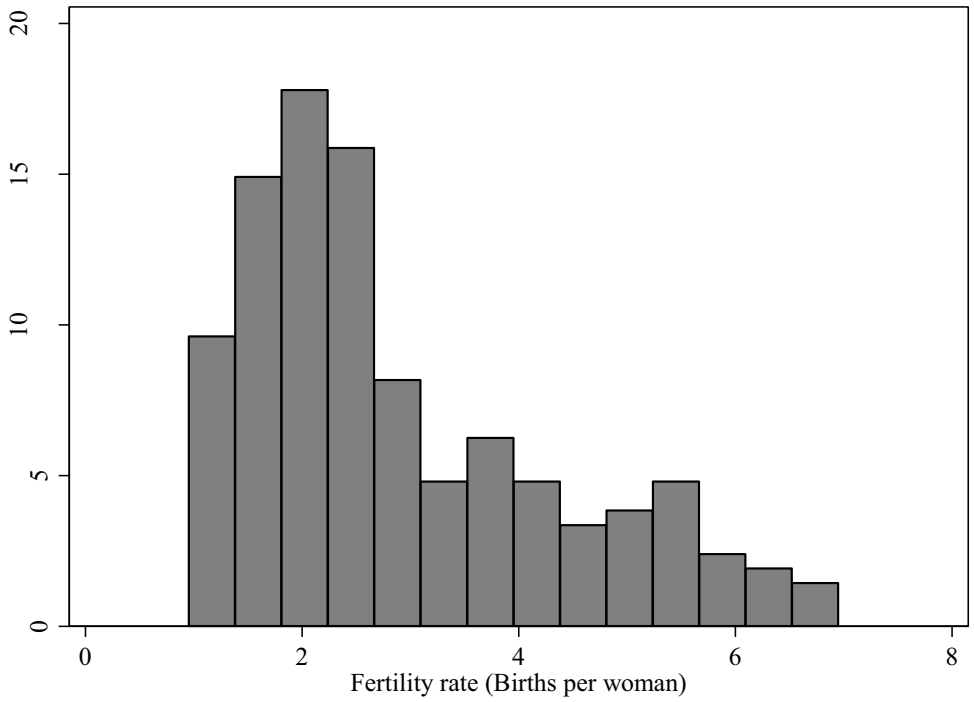

Figure 3. The International Distribution of Fertility in 2007.
Data source: World Development Indicators 2009. *Notes*: (a) The fertility rate is defined as the number of children a woman will give birth to if she lives through her reproductive life span. (b) 208 countries are represented in the sample. (c) The lower decile is 1.4 births, the lower quartile is 1.8 births, the median is 2.4 births, the upper quartile is 3.8 births, and the top decile is 5.3 births.

Hence, if we ignore deaths the figure would tell us that population growth varies from about one percent to roughly five percent. Naturally, this exaggerates global differences in population growth since high fertility regions usually are also characterized by high mortality, while low fertility regions are characterized by low mortality. In 2007 population growth therefore "only" varied from about -0.15% to 3.1%.[13] How much does this kind of variation matter to the evolution of productivity?

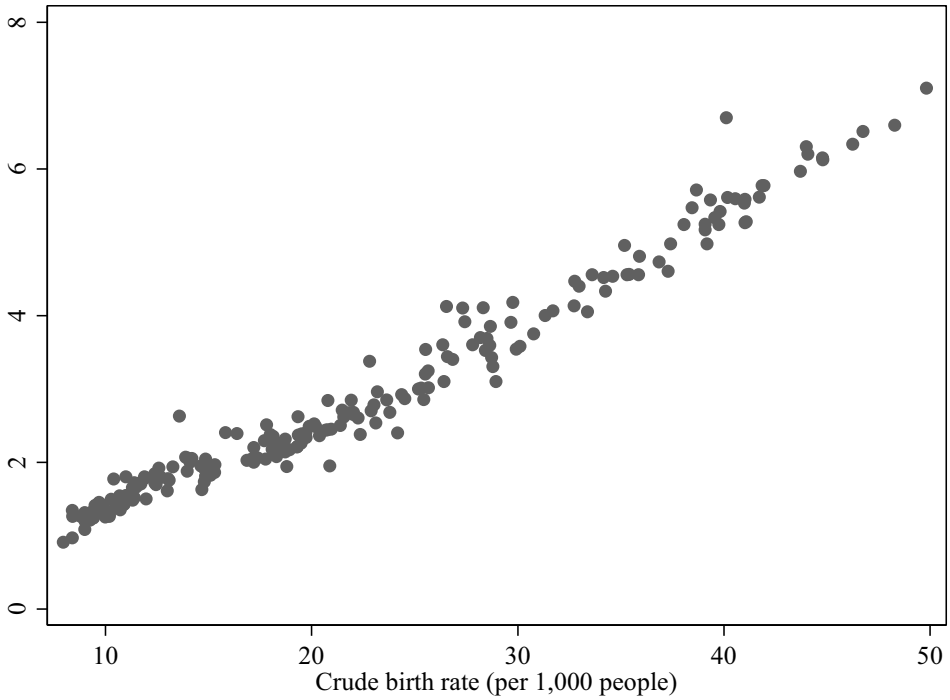

Figure 4. Fertility vs. Crude Birth Rate in 2007.
Data source: World Development Indicators 2009. *Notes*: (a) The fertility rate is defined as the number of children a woman will give birth to if she lives through her reproductive life span. The crude birth rate is the total number of births per 1000 people. (b) 208 countries are represented in the sample.

The answer to this question depends on the importance of land. That is, loosely speaking, how important agriculture is. The reason is that the "capital dilution" effect, mentioned above, is very powerful if land is important to production. If land is important, a larger population size works to reduce the level of labor productivity; larger population growth therefore works to lower labor productivity *growth*.

[13] This is a comparison of the 5th percentile (low population growth) to the 95th percentile (high fertility) across the 208 countries depicted in Figure 3.

In practice there is good reason to believe land is important in high fertility regions. Figure 5 shows the relation, across 149 countries, between the size of the agricultural sector (% of GDP) and the Crude Birth Rate; the correlation coefficient is 0.70.

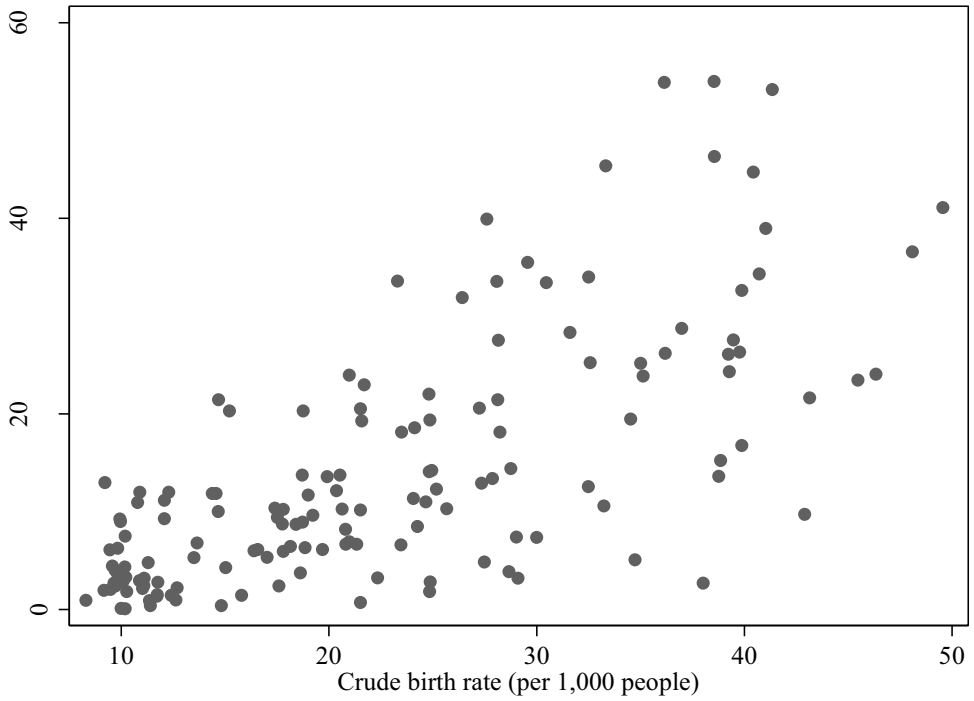

Figure 5. Crude Birth Rate vs. the size of the agricultural sector in 2007.
Data source: World Development Indicators 2009. *Notes*: (a) The crude birth rate is the number of births per year per 1000 people. (b) 149 countries are represented in the sample.

If land therefore is an important input, population growth will provide an important drag on growth. As a rule of thumb, the growth rate of labor productivity will in this case decline with population growth by an amount that depends on the ratio between land's share in GDP and capital's total share in GDP. [14] In national accounts "capital's share" is about 0.4; but it convolutes both land and other forms of physical capital (like equipment and buildings). Hence, there is no simple way of gauging the importance of land in output. A conventional estimate, however, would be a land share of about 0.1. If so, then a reduction in population growth by three percentage points per year will increase labor productivity growth by about 0.75 percentage points per year, given a total capital share of 0.4. If the reduction is persistent this will have a non-negligible impact on living standards. An increase

[14] See Annex 1 for the calculation.

in growth by 0.75 percentage points means that livings standards 25 years later are 20% higher than what they would have been if population growth did not decline. If, on top, declining fertility is associated with better health and more schooling, the total impact on living standard will be larger still. The scope for per capita income gains through fertility reduction is therefore quite respectable *a priori*.

The problem is how to bring down fertility in practice, by way of policy. Family planning programs may help. But it is important to recognize that the observed variation in fertility rates across countries is unlikely to be caused by "unwanted" births. Actually (but not surprisingly), people around the globe tend to have about as many children as they want to have (Prichett, 1994). Hence to reduce fertility markedly and persistently *desired* fertility needs to decline. Obtaining reductions in desired fertility is not easy, as it depend in complex ways on institutions, culture, mortality and the value of having an education. In reality, therefore, it seems unlikely that fertility is directly amendable to policies such as foreign aid disbursements. Naturally, there are exceptions; China's "one child policy", for instance. But in many cases policy makers may well be weary of following this rather intrusive approach.

These comments are associated with an important caveat. Perhaps surprisingly, no aggregate study seems to exist that tries to evaluate whether aid efforts have had a measureable impact on fertility. Hence, while we are somewhat sceptical that family planning efforts have had a major impact on fertility, for reasons noted above, it is admittedly an open empirical question.

2.3. Channel 3: Productivity

A key stylized fact in comparative development is that the bulk of cross-country differences in labor productivity can be accounted for by this last channel: productivity (Caselli, 2006). While this is an important piece of information, it is also (at first sight) a little discouraging. In practice "productivity" is simply the left-over residual, when output per worker is "pruned" from the influence from physical capital and human resources. In this sense, it is a measure of ignorance as much as a measure of productivity. Still, on *a priori* grounds a fairly short list of factors should be key proximate determinants of the residual.

For starters, the level of technological sophistication of an economy should influence over-all productivity. This suggests that efforts aimed at facilitating the adoption and invention of new techniques would be important. Providing a reasonable quantitative assessment of the

strength of this channel is still difficult. Hence no additional space will be dedicated to the issue here. Fortunately there are alternative channels that are more amendable to scrutiny.[15]

In recent years, considerable effort has been extended in the examination of two complementary, and perhaps more concrete, determinants of aggregate productivity: aggregate efficiency and public goods provision.

The efficiency at which scarce input factors, like physical capital and human resources, is distributed across occupations, industries and sectors ought to leave an imprint on productivity; the higher the level of efficiency, the greater the level of productivity. More broadly this channel speaks to the *interaction* of individual sectors of the economy which together makes up total output, or GDP. Public good provision (chiefly in the sense of infrastructure) will also work to elevate over-all production for total physical capital and human resource endowment given. As we shall see these two separate elements interact in important ways.

Consider the issue of economic "infrastructure". An economy's infrastructure has many dimensions ranging from traditional infrastructure such as railways, roads and harbors, to more contemporary notions involving the electricity infrastructure, and the "virtual" infrastructure in the sense of the Internet. All of these elements are important to economic activity. As a result, it is no surprise that they all are highly correlated with per capita GDP. Figure 6 illustrates the association between one dimension of the infrastructure, the telephone network, and GDP per capita. It is visually obvious that the two variables are highly and positively related, the correlation coefficient is 0.84.

[15] Besides, the "pure" technology channel may be less important than the two alternative avenues of influence of productivity to which we now turn. In particular, Chanda and Dalgaard (2008) show that up to 85% of the comparative differences in productivity across countries emanates from *relative* efficiency across sectors (agriculture vs. industry and servies).

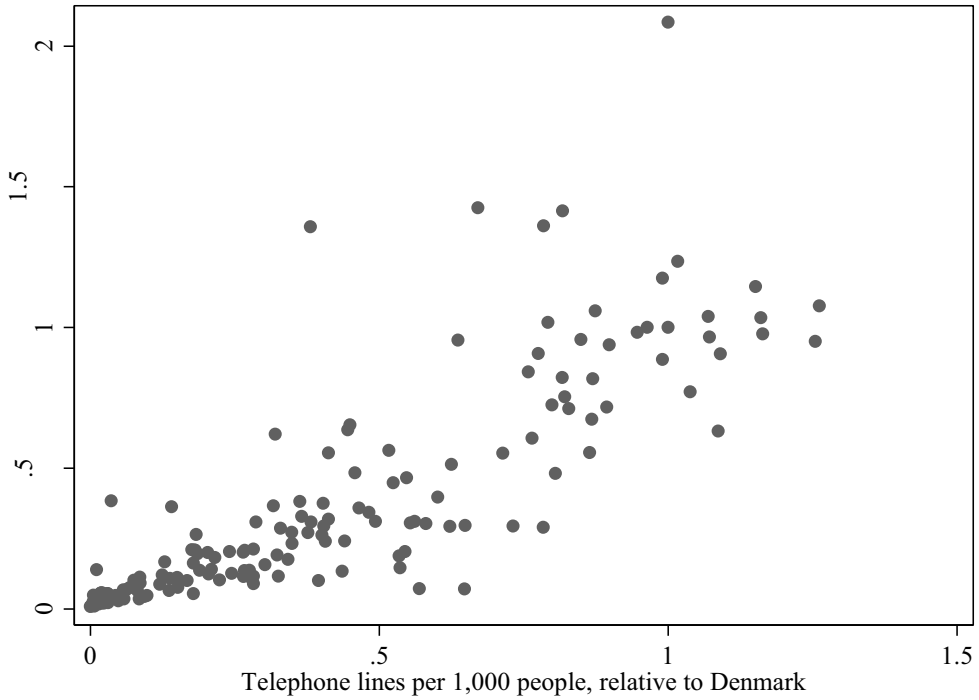

Figure 6. The association between PPP adjusted GDP per capita and Telephone mainlines per 1,000 people capita in 2007 (Relative to Denmark).
Data source: World Development Indicators, 2009.

Naturally, correlation does not prove causation. A part of the above correlation is unquestionably due to the fact that richer nations invest more in their infrastructure than poor countries do. Sorting out the relative strength of these two separate chains of causation is therefore generally very complicated.

But in the context of telecommunication careful studies exist which try to pin down its impact on aggregate productivity. Röller and Waverman (2001) estimate for a sample of OECD countries that raising *the penetration rate* of telecommunication (defined as mainlines per capita) by one percent increases productivity by 0.15 percent.

The size of the effect is elevated once it interacts with the "capital multiplier" discussed above. In the long-run, therefore, an increase in the penetration rate of one percent can be expected to increase GDP per capita by about 0.23 percent, in a closed economy setting.[16]

[16] Annex 2 contains some technical details on this point, for the mathematically inclined reader.

How much potential does this channel hold in terms of fostering growth? As it turns out, the amount of variation in telecommunication infrastructure across the world is huge. In fact, going from the places with least infrastructure (a penetration rate of about 0.004) to the places with the most infrastructure (a penetration rate of about 0.93) yields an astonishing difference of 92.6 percentage points. To put it differently, if the penetration rate in the poorest places were increased to the level of the richest places this could increase income per capita by around 20%.[17]

Yet, the impact may well be larger still, due to the efficiency channel mentioned above. The magnification arises due to *linkage effects*, which produces a potentially powerful additional multiplier. If productivity in a given sector expands it will affect the productivity of other sectors that use its output as input, and so on. Like ripples in a pond, increasing productivity in one sector may spread to the rest of the economy. While the notion that "linkage effects" are important is an old idea in development economics, recent work by Jones (2009) has re-vitalized the idea by developing a formal framework that allows one to gauge the size of the "input-output" multiplier in a transparent way. In the present context the input-output multiplier plausibly doubles the impact from expanding the penetration rate; from about 20% to around 45%.[18] Hence, the impact from just telecommunication on productivity easily towers that of the other channels discussed so far.

While these calculations are encouraging they come with an important warning. The estimate of telecommunication's impact on productivity, which these calibrations are based on, pertains to a set of OECD countries; the impact could be smaller in poor places.

A particular reason why this may be the case is that telecommunication likely is complementary to other forms of infrastructure, which we have *not* taken into account. This point is illustrated in Figure 7, which shows the association between per capita GDP and a number of additional dimensions of infrastructure. As is clear from the figure; countries lacking in roads, are also poorly endowed in term of railways, electricity infrastructure, and so on.

This is important as these separate "imperfections", according to Jones (2009) analysis, may influence the size of the input-output multiplier greatly. The basic intuition is that investments in (say) telecommunications at the firm level may have a small impact on productivity if there are no roads by which the firm may transport its goods. Similarly, while

[17] The calculation is $(0.93/0.4)^{0.23}$, which comes to about 1.21, or roughly 20%.

[18] Annex 3 contains some details.

phone lines may be central to the expansion of Internet connectivity, the presence of a well functioning power supply is equally important. Without the latter, the former will have a small effect on ICT diffusion. Hence, a large impact from telecommunication, as estimated by Röller and Waverman, most likely needs to be assessed in light of the fact that many other dimensions of the infrastructure "works" in rich places.

These calculations pertain to a closed economy. How the analysis would change if an open economy setting is considered is hard to assess quantitatively. But a brief discussion may make clear what is involved.

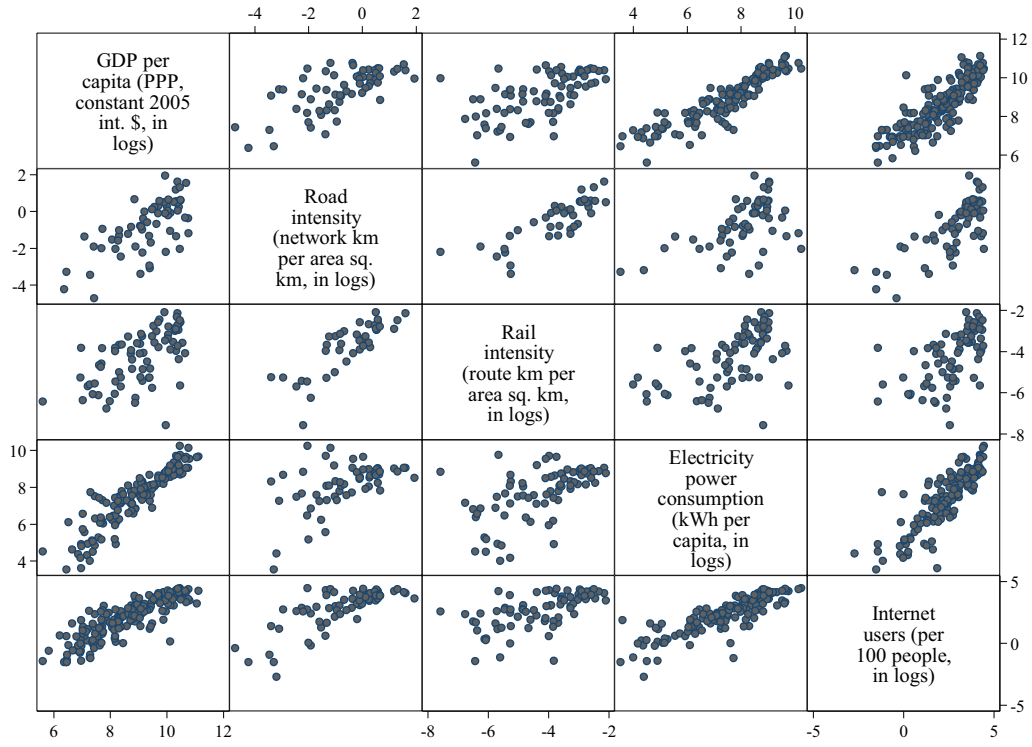

Figure 7. The relations between GDP per capita, Roads, Railways, Electric Power consumption, and Internet users in 2005.
Data source: World Development Indicators 2009.

First, allowing for capital mobility is unlikely to reduce the impact from infrastructure much. The reason is that infrastructure investments expand productivity, which raises the return on capital investment. Hence, rather than crowding private capital out, it likely crowds private capita in. Second, however, the open economy may also influence the input-output multiplier. It could increase it, if the economy in question suffers from many imperfections

that are "done away with" by the (implicit) introduction of foreign firms. Still, mechanically, the input-output multiplier could also be diminished if foreign firms "sever" the links between domestic firms; along the lines of how import tends to reduce the Keynesian multiplier in introductory economics textbook treatments of the short-run determination of output. In the end, the net impact of allowing for an open economy setting on the strength of the infrastructure-productivity nexus is somewhat unclear at the moment.

In sum, it seems reasonable to assess that the scope for improvements in productivity, by way of infrastructure investments, is substantial. But that is not to say that it constitutes a "silver bullet" as each separate element in an economy's infrastructure likely interacts in important ways. Fixing one "ailment" may not cure the patient. In addition, the economic costs involved in the above types of investments are likely very considerable.

2.4. Summary

For every percentage point investment in **physical capital** the impact on the long-run *level* of productivity will be about 0.5 – 0.67 %, if the economy is *closed* to global capital flows. If the economy is *open* to capital mobility, capital investments are unlikely to impact on labor productivity.

Measures aimed at **health** can be expected to raise productivity, but only in the long run. Plausible detection lags are of the order of magnitude of four to six decades, in the closed economy scenario. As discussed above, if health interventions lower mortality rates in a *closed* economy scenario GDP per worker can be expected to decline initially, due to increasing relative scarcity of capital and land. Only eventually will the productivity enhancing effect of increased "vitality" of the labor force manifest itself. The prospects are somewhat better if the economy is *open* to international capital flows, as international capital imports may off-set the tendency for physical capital to be "diluted" from an increasing labor force. The magnitude of the impact from health interventions in the aggregate is difficult to assess, and undoubtedly case specific (i.e., depends on which diseases are targeted, the initial disease burden, etc.). Nevertheless, one should be realistic about the total impact. In the case analyzed above a major intervention which succeeds in raising life expectancy by 20 years, plausibly leads to increases in the *level* of labor productivity from about 15% (closed economy) to about 25% (open economy).

The prior for **schooling** would involve similar though somewhat larger magnitudes: one additional year of schooling on average plausibly increases labor productivity by about 10% provided schooling does not affect health or fertility.

Targeting **fertility** is in many ways more promising from the perspective of potential impact. For each percentage point of fertility reduction per capita GDP *growth* will likely increase by 0.25%. As explained above, the size of this effect is likely to decline, however, as agriculture becomes less important to economic activity. A difficulty, however, lies in how to reduce desired fertility.

Finally, **productivity** is a key factor one may try to target. As explained above, in less developed economies the impact may be substantial; a possible impact range is that each percentage point of GDP invested in infrastructure may increase the level of long-run productivity by between 0.25% and 0.45%. The impact is likely to be intertwined with the stage of development for all the economically relevant dimensions of infrastructure.

The overall conclusions, which we will carry forward into the discussion of the evidence is that :

(i) A reasonable estimate for the elasticity of aid, when measured in terms of GDP or GNP, with respect to long-run labor productivity is somewhere in between 0.25 and 0.7. That is, increasing the aid-to-GDP ratio by one percent leads to an increase in the long run level of per capita GDP of 0.25% to 0.7%. This summarizes *best-case* elasticities if aid is targeted towards capital and infrastructure, which constitutes a considerable part of total aid interventions in practice.

(ii) The impact from aid likely varies across countries. The discussion above has motivated that the impact from aid to physical and human capital accumulation varies depending on whether the country in question is open or closed to international capital movements. The aggregate effect of a fertility intervention would depend on lands' share in GDP (the size of agriculture), and thus be country specific. Investments in particular forms of infrastructure will depend on all other forms of infrastructure and their efficiency. As explained in Dalgaard and Hansen (2009) there are many other reasons why the impact may be country specific beyond these.

3. Aid Effectiveness: The evidence

Dalgaard and Hansen (2009) discusses the general methodology involved in assessing aid effectiveness empirically, namely regression analysis. The evaluation study also lays out the key methodological stumbling block to this endeavor: the identification problem. The fundamental problem is that aid is both a likely cause and a consequence of growth (via the allocation of aid). As a result, estimates obtained using the standard methodology ("ordinary least squares") will be misleading. Instead, in order to elicit information about the causal impact of aid on growth a viable approach is instrumental variables estimation.[19]

To employ this estimation procedure, the critical step is to obtain a viable instrument for aid, which has the following properties: (i) it explains part of foreign aid inflows to the recipient, (ii) it is not affected by economic growth, and (iii) it *only* affects growth *through* the impact of foreign aid. With an instrument in hand, that fulfills (i)-(iii), it is possible (in principle) to disentangle the impact of aid on growth from the impact of growth on aid.

The discussion to follow is structured according to the nature of the instrument invoked. Essentially, the literature on aid effectiveness has proceeded along three broad tracks since Burnside and Dollar (2000) and Hansen and Tarp (2001), which are the two studies that largely ignited the recent surge of academic interest in aid effectiveness.

The first set of contributions essentially followed Hansen and Tarp. That is, they examine the impact of total aid on growth, using what is called "*internal*" instruments. Internal instruments (loosely speaking) refers to the practice of employing lagged values of aid to identify the impact of aid on growth. We review this literature first.

The second track involves a the use of "*external*" instruments, and can be seen as methodologically following Burnside and Dollar. In these studies identification does not rely on lagged aid being useful as an instrument. Instead the researcher introduces an instrument for aid, which in no form enters the growth regression. This is the second set of contributions that we review.

Finally, a set of contributions have tried to move beyond the study of *total* aid, and begun to analyze the impact of particular forms of aid: aid modalities. This is the third set of studies that we review in this section.

---

[19] See Dalgaard and Hansen (2009), Section 4, for details on the general methodology.

3.1. Reduced form evidence: The impact of total aid receipts

3.1.1 Internal instruments

In 2001 Hansen and Tarp (2001) published the first implementation (in the literature on aid effectiveness) of a solution to the identification problem which basically involves using past realizations of foreign aid flows as an instrument for current foreign aid.[20]

The basic logic of this approach is that past aid allocations in practice is a fairly good guide to current aid allocations; aid donations are rather persistent over time. That is, previous realizations of foreign aid has strong explanatory power vis-a-vis current aid inflows. This is the first requirement for a viable instrument. Second, aid allocations in the past cannot be affected by current growth. Hence, past aid levels also fulfill the second requirement listed above. Third, it may seem reasonable at first to suspect that past aid levels does not matter to growth in the current period, conditional on current aid flows.

The key result in Hansen and Tarp (2001) is that foreign aid does increase growth, but is subject to "diminishing returns". That is, the marginal impact of aid on growth declines as the total inflow increases. This result suggests the impact of foreign aid is nonlinear (see Dalgaard and Hansen, 2009 Section 6); the effect of foreign aid will be larger in countries with more modest inflows of aid. Following Hansen and Tarp (2001) a considerable number of studies have appeared which employ the same approach to identifying the impact of aid on growth. The difference to the Hansen and Tarp contribution mainly lie in what sort of non-linearity the individual authors argue in favor of.

In a recent survey Roodman (2007) takes stock, providing rigorous robustness checks of the most influential contributions in this area. The somewhat discouraging result is that essentially all the studies that Roodman survey are fragile in one way or another. The only exception seems to be Dalgaard, Hansen and Tarp (2004) (DHT), which passes essentially all checks. In the remaining part of this section we therefore focus on this contribution.

The key finding in DHT is that the impact of aid varies according to a climate gradient; from the 1970s until the mid 1990s the impact of aid was smaller in countries situated in tropical climate zones. It is natural to wonder what explains this finding. Since it is hard to believe that the climate matters *per se* the cause of the non-linearity is likely to be an indirect one.

---

[20] Formally, the method is called GMM estimation of dynamic panel data models (Arellano and Bond, 1991). Hansen and Tarp (2001) themselves were reacting to a paper by Burnside and Dollar (2000) which can be seen as the epicenter for the recent literature on aid effectiveness. We return to that contribution in the next section, since they employ a different approach to identifying the impact of aid on growth.

As explained in Section 2 the impact of foreign aid is likely to depend in various ways on the structural characteristics of the economy receiving the flows. Hence, it is likely that "tropical area" simply is capturing such characteristic, which matter to aid effectiveness. Investments in human resources may serve as an illustration.

Tropical areas tend to face unique challenges in regard to diseases, which influences health outcomes. As a matter of stylized fact: in 2000 average life expectancy was on average 64 in countries where more than 50% of the land area is situated in the tropics; in countries with less than 50% of their area in the tropics the comparable number was nearly ten years higher, at 73. If indeed health outcomes are poorer in tropical areas this can motivate a lower impact from aid, since the effect of human capital investments may decline because learning outcomes for children recede with poor health conditions. As a result, aid funded investments in schooling may lead to less growth in tropical areas.

Tropical areas are also distinguishable in terms of fertility. As a matter of stylized fact: whereas the demographic transition began (on average) in 1948 in countries with most of its land area outside the tropics, the comparable year for countries with the bulk of their area located inside the tropics is 1980. This difference can motivate a differential response to aid transfers vis-à-vis fertility, which in turn is important to growth as explained in Section 2. The reason is that prior to the demographic transition increases in household income usually translates into larger family sizes; after the demographic transition improvements in average income tends to translate into intensified investments in schooling and smaller family sizes. It is therefore possible that aid transfers may have stimulated population growth in tropical areas, and reduced it outside the tropics, thus accounting for the variation in aid effectiveness we find in the data.

As discussed in DHT, other explanations for the difference in impact from aid inside and outside the tropics may be constructed. But naturally, the study cannot pinpoint what the reason really *is,* which is a weakness of the approach.[21]

Leaving this difficulty aside and turning to the results, the study by DHT supports an elasticity of about 0.4 outside the tropics (the best case scenario). That is, increasing the aid-to-GDP ratio of 1 percent increases the long-run level of GDP per capita of about 0.4%. In light of the discussion in Section 2 this estimate must be viewed as reasonable. At the same time it should be recognized, that the impact of aid inside the tropics is in fact estimated to

[21] The reason why the study did not use interactions with life expectancy (for instance) is that these variables themselves tend to be endogenous. This *elevates* the identification problem further (see Dalgaard and Hansen, 2009).

be *negative*. This could be consistent with a capital diluting effect of population growth, the latter of which may be spurred by aid inflows.

Ultimately the plausibility of these estimates hinges on more than just their absolute value and interpretation. The identification strategy is central: if conditions (i)-(iii) listed above are not fulfilled the results are suspect.

Unfortunately, in the end it is reasonable to be suspicious of the approach. As highlighted above the approach involves the use of past aid as a means to obtain identification; past aid serve as an instrument for current aid. This implies that past aid *must* be irrelevant to current growth, conditional on current aid. If we think about aid investments in physical capital (equipment), or of aid as a transfer of income to poor households, this assumption may be plausible. But if aid funds investments in humans (say, schooling or health) or infrastructure (e.g., roads) it is far from obvious that the assumption is a good one. The reason is that there is a considerable production lag involved in producing human capital and infrastructure; investments undertaken today only holds productive benefits years into the future when the educated labor force enters into employment, or when the road is opened. Since parts of total aid flows are used for such purposes the crucial exclusion restriction mentioned above is suspect on theoretical grounds.

Formally, the exclusion restriction can be submitted to statistical testing.[22] Unfortunately, the kinds of test invoked are not very informative. Technically speaking what is tested is the statement that the exclusion restriction *is* fulfilled. Under "the best of circumstances" (from the point of view of the validity of the study) a researcher may *fail* to *reject* this hypothesis. But a failure to reject is not the same as saying the hypothesis is true. In fact, the exclusion restriction may not be fulfilled and yet the researcher fails to reject it. Consequently, few researchers are comfortable relying on these sorts of tests absent a very strong prior that the exclusion restriction is plausible. Admittedly, this is not the case here.

It is worth observing that this is a general problem with *every* study that employs "internal instruments" in an effort to identify the impact of total aid on growth. At this stage a reasonable conclusion is that the practice of using internal instruments in an effort to obtain information about the impact of total aid on growth is unlikely to lead to firm conclusions. Instead one may see two ways forward.

---

[22] A formal test requires "overidentification". That is, you need more than one instrument for one "endogenous variable" (in this case aid). In practice the GMM estimator allows one to use several lags of aid, and lagged differences in aid. Hence, overidentification is always achieved in these studies.

The first route would be to obtain identification using external instruments. That is, an approach which does not rely on lagged aid in the estimation procedure. The second route might involve disaggregating foreign aid. As alluded to in the discussion above, there might be parts of total aid flows for which lagged aid is plausibly excludable, but external instruments with a firm theoretical motivation is clearly to be preferred here as well. We turn to these two approaches next.

3.1.2. External instruments

Burnside and Dollar (2000) (BD) is the paper which started something of a land slide of work on aggregate aid effectiveness. The key result claimed by the authors is that foreign aid seems to have a larger (i.e., positive) impact on growth in countries that pursue sufficiently "good policies"; BD define a "policy variable" which nests inflation, budget balance and openness to trade.

BD, like (most of) the ensuing literature, recognize that aid flows undoubtedly are endogenous; affected by growth (or its determinants) as well as affecting growth. In order to deal with this problem the authors proposed a set of variables that must fulfill the three criteria for valid instruments listed above.

Their instruments of choice include (but are not limited to) population size, imports and exports of arms, and population interacted with past realizations of the policy variable. While population size and arms trade are clearly external instruments (i.e., they in no way show up as growth determinants in the regression), the interaction effect involving lagged policy is really an "internal" instrument. Hence the BD approach is strictly speaking somewhere in between external and internal instruments.

Subsequent research has found BD's results to be fragile, and to not hold up in out-of-sample tests (e.g., Dalgaard and Hansen, 2001; Easterly, Levine and Roodman, 2004; Roodman, 2007). To this one may add that the chosen instruments are not particularly convincing. As argued by Rajan and Subramanian (2008), there is no particular reason why past policy choices should be irrelevant to current growth. Structural policy choices should (at least) be expected to have a long reach. Policy irrelevance, outside the current period, is however required for BD's results to be valid. Hence, much like the contributions discussed in the last section the BD list of instruments are theoretically unappealing.

Thus, a challenge for subsequent research has been to improve upon the identification strategy. In this respect Rajan and Subramanian (2008) (RS) is arguably the most innovative

study on aid effectiveness. The key innovation by RS is the attempt at constructing an *a priori* plausible instrument for aid. The construction of the instrument proceeds in a two steps.

First, the authors explore the *supply* of aid by estimating bi-lateral aid flows between donor and individual recipients. The chosen specification is inspired by the aid allocation literature, which for years have examined the question of why some countries receive more aid than others, while being (surprisingly) detached from the effectiveness literature. More specifically, the authors show that three broad categories of variables hold considerable explanatory power vis-à-vis aid flows: (i) initial (1960 *viz.* 1970) relative population size (donor size/recipient size), (ii) common language and (iii) past and present colonial ties. The logic of the first variable is that it captures donor influence; donors are, the argument goes, more likely to aid countries that they can exert more influence on. Donations should therefore be increasing in relative donor size capturing the strategic motive for aid donations. Note that this element is similar to BD, who also uses population size (of the recipient) as an instrument for aid. The logic of the two other determinants is that they capture historical ties. Donors with a historical (or linguistic) affiliation with a less developed country are more likely to aid that country. A useful characteristic of these variables is that they all are plausibly exogenous to growth in the sense that growth from 1960 to 2000 (viz 1970-2000) cannot *affect* their values.

The second step is to derive the amount of aid predicted by these variables for each recipient, by summing across donors. The end result is a variable which includes aid allocated to any recipient for reasons *unrelated* to growth; this is their instrument for aid. Thus armed, the authors proceed to estimate the impact of aid by way of cross section analysis.

Table 1 reports key results from the RS study. Column 1 and 2 show the basic ordinary least squares (OLS) regression results for the periods 1960-2000 and 1970-2000, respectively. As explained in Dalgaard and Hansen (2009) these sorts of results are very unlikely to reflect the impact of aid on growth. Instead they merely constitute a (partial) correlation between growth and aid; OLS results are likely to yield a negative association between aid and growth, reflecting that donors prioritize slow growing counties. This is indeed what RS find: the coefficient on aid is negative and statistically significant.

Table 1: Aid effectiveness revisited

Dependent variable: Growth in GDP per capita						
	I	II	III	IV	V	VI
Aid	-.063**	-.07*	.06	.09	.07	.10*
	(.02)	(.04)	(.05)	(.06)	(.05)	(.06)
Initial income	-1.33***	-1.67***	-1.17***	-1.40***	-1.15***	-1.42***
	(.28)	(.31)	(.34)	(.39)	(.34)	(.37)
Trade openness	1.78***	2.27***	1.62***	2.13***	1.8***	2.24***
	(.43)	(.47)	(.59)	(.56)	(.50)	(.53)
Life expectancy	.02	.01	.05**	.075**	.04*	.06**
	(.02)	(.03)	(.02)	(.03)	(.02)	(.02)
Geography	.34**	.38**	.52***	.60**	.49**	.48***
	(.13)	(.17)	(.16)	(.23)	(.15)	(.18)
Institutions	3.94***	4.02*	4.55***	4.07**	5.91***	5.57***
	(1.49)	(2.22)	(1.52)	(2.10)	(1.39)	(1.68)
Inflation	-.002	-.003	-.002	-.005		
	(.002)	(.003)	(.003)	(.004)		
Real money balances	.01	.016	.016	.01		
	(.01)	(.01)	(.011)	(018)		
Budget balance	-.007	-.01	.015	.01		
	(.02)	(.03)	(025)	(.03)		
Revolutions	-1.26***	-1.3***	-1.14**	-1.40**	-.87	-1.08*
	(.50)	(.48)	(.55)	(.59)	(.63)	(.62)
Ethnic Fraction.	-.10	-.39	.71	.78		
	(.44)	(.70)	(.54)	(.77)		
Sub-Saharan Africa	-.74	-1.1*	-1.33***	-1.67***	-1.46***	-1.79***
	(.45)	(.59)	(.43)	(.56)	(.38)	(.54)
East Asia	.45	.52	.55	.57		
	(.32)	(.41)	(.40)	(.45)		
Period	1960-2000	1970-2000	1960-2000	1970-2000	1960-2000	1970-2000
Estimator	OLS	OLS	2SLS	2SLS	2SLS	2SLS
N	74	78	74	78	74	78

Notes: All regressions contain a constant. Robust standard deviations in parenthesis. *,**,*** denotes significance at 10,5 and 1%, respectively. For data description, see Rajan and Subramanian (2008).

Columns 3 and 4 reproduce the results for the two sub-periods, when the instrument is employed. This is the so-called two-stage least squares estimates (2SLS for short). The key result advertised in the paper is that, once instrumented, the link between aid and growth is no longer significant. This can also been seen from the table; at conventional levels of significance one is unable to reject a hypothesis involving a zero impact from aid on growth.

The insignificance of aid none withstanding, it is interesting to observe that the point estimate changes rather radically from columns 1 and 2, to columns 3 and 4. The point estimate of aid more than doubles, going from negative to positive. This is precisely what one would expect if the OLS coefficients are capturing mostly that donors are giving aid to slow growing countries, rather than reflecting a potential positive impact on growth.

In terms of the economic impact the result also seems believable; evaluating the elasticity of aid with respect to the level income we find an elasticity about 0.3. This result fits reasonably well with the bounds laid out in Section 2.[23] Still, we cannot reject the hypothesis of a zero impact from aid on growth. This implies that we cannot reject that aid, on average, has been ineffective in raising living standards. While the RS study arguably is the most convincing around, nothing is unfortunately foolproof. There are two sets of concerns worth emphasizing.

The first concern is that the key result is a little too fragile for comfort. Consider again Table 1, column 3 and 4. Comparing the results to column 1 and 2 it is clear that using instruments dramatically increases the impact from aid judged from the point estimate. Hence, from a technical perspective, the main reason why aid is insignificant is that the 2SLS estimate (column 3 and 4) is estimated with greater imprecision; the reported standard deviations in column 3 and 4 are about double those of column 1 and 2.

A reason why the standard deviations might be "inflated" is that aid is correlated with other variables included in the regression.[24] Reexamining the RS specification one may note that several of the variables are insignificantly correlated to growth. The list includes various

---

[23] For the technically interested reader, the calculation proceeds as follows. In RS sample the average aid-to-GDP ratio is 6.1 (percent). The coefficient reported in Table 1 for aid 1960-2000 is 0.06. The regression also includes a control for GDP per capita, which is significant and enters with a coefficient of -1.17, capturing convergence. In the long-run growth is plausibly independent of aid (given, instead, by technological progress); then the elasticity of the level of income is 0.06*6.1/1.17, which is about 0.3.

[24] That is, so-called multicollinarity affects the precision of regression estimates. In addition, however, it is also quite normal for 2SLS estimates to be more imprecisely estimated than OLS estimates since they inherently involves two, rather than one, regression: one fitting aid to the instrument, and one involving fitting instrumented aid to growth.

policy variables (inflation, real money balances and the budget deficit), an East Asian dummy and a variable which captures ethno-linguistic fractionalization (ELF).[25] As it turns out, these variables are highly correlated with aid. Collectively, they motivate (not shown) some 30% of the variation *in aid* during the 1970-2000 period, suggesting that while uncorrelated with growth (conditional on other included variables) they may in fact be inflating the estimated standard deviation of the aid parameter due to colinearity. The degree of collinearity between aid and the controls is smaller during the period 1960-2000, but still fairly substantial.

Moreover, on theoretical grounds, there is actually little reason why the insignificant controls should be in the regression to begin with. For instance, the insignificance of the policy variables is rather unsurprising since RS also includes a measure of institutions in the regression, which encompasses the effectiveness of bureaucracy; undoubtedly a determinant of policy. There is more reason to believe that ethno-linguistic fractionalization (ELF) could matter structurally. But since the regression also involves a variable capturing revolutions (thus, a measure of political instability and conflict), it is less surprising that ELF is insignificant, and can be omitted since these variables captures the main avenue through which ELF should matter to growth. The insignificance of the East Asian dummy simply suggests that the Asian growth experience is well captured with the variables otherwise included in the regression. Against this background it is not disturbing to find that *all* of these variables can be simultaneously excluded from the growth regression at conventional levels of confidence. [26] Hence on statistical as well as on economic grounds one can feel comfortable omitting these variables from the analysis.

Column 5 and 6 of Table 1 reveal how the results of RS are modified if these variables are omitted. For both periods the parameter uncertainty is reduced judged from the standard deviations, as expected. The effect is stronger for the 1970-2000 period in keeping with the observation that the insignificant growth determinants are more highly correlated with aid during this time interval. Insignificance for the 1960-2000 period is not overturned, but it is for 1970-2000. For the period 1970-2000 we reject a zero impact hypothesis with more than 90% confidence. The point estimate for aid is essentially unaffected, and the remaining variables are significant, and of the right sign. This demonstrates the fragility of the "insignificance result" claimed by RS.

---

[25] The " Asian dummy variable" is an indicator taking on a value of "1" if a country is situated in this particular region.

[26] F-tests of the joint exclusion of all insignificant variables in Columns 3 and 4 return *p*-values of 0.21 and 0.18.

The second concern is whether in fact the instrument employed is excludable. One of the key requirements for an instrument to be valid is that it is does not affect growth once we have taken aid into account. This means - in the present context - that country size (measured by population), the spoken language and past colonial ties *cannot* matter to growth, beyond via aid donations. This is not necessarily a very believable requirement.

For one thing, it is plausible that larger countries (all else equal) have an advantage in the growth process by way of larger markets, which facilitates division of labor and the exploitation of increasing returns to scale. If so, then the instrument employed for aid is likely invalid, as it encompasses local population size. Similarly, past colonial ties may well have affected institutional developments, including how the legal system operates. If true, then the information of colonial ties will matter to growth, beyond its impact on aid. Language is a conveyer of cultural values, which might relate to a host of growth determinants: how much individuals choose to save, preferred fertility, gender roles, societal trust, and so on. In the end one is left with the strong impression that a key identifying assumption of zero impact from population, colonial ties and language on growth (conditional on aid), is questionable. With only one instrument there is no way one can formally test this assumption is infeasible; it boils down to a matter of faith.[27]

The study by RS is unquestionably innovative in the context of aid effectiveness research. Nevertheless, it is not beyond critique; the critical exclusion restriction can be questioned as argued above. Moreover, the key result – that aid has been ineffective – is not robust.

In addition, one may wonder whether it is reasonable to assume that the effect of aid is linear.[28] Indeed, most of the literature between Burnside and Dollar and RS provides evidence of some form of nonlinearity. Moreover, as discussed in Dalgaard and Hansen

---

[27]The authors are aware of this difficulty and try to strengthen the case by examining a *second* instrument, land area in km^2, so as to facilitate a formal test of excludability of instruments. Still, they are clearly ambivalent about this second instrument, observing (footnote 16):" *While a measure of country size could in itself be a plausible instrument, the reason not to make it the preferred one is that there is uncertainty whether it can satisfy the exclusion restriction; that is, a number of reasons can be advanced as to why a recipient's size would have an independent effect on growth.*". The problem is however that country size already is in the instrument via *population size*. The fact that the instrument also contains information about language and colonial ties does not eliminate this problem. In the absence of a strong prior that *all* instruments are excludable, the tests used to check excludability have notoriously low power and may thus be misleading.

[28] RS do in fact examine various nonlinearities, but claim them to be insignificant. A difficulty with most of these exercises is that RS only has one instrument; as explained in Dalgaard and Hansen (2009) the endogeneity of the interaction term should not be ignored. In one case, where RS explores the nonlinearity purposed by Dalgaard et al. (2004) no additional instrument is required as country area in the tropics is exogenous. However, RS do not exactly use country area in the tropics, but rather a variable where tropics is combined with separate information about rainfall. In other instances the "interacting" variable is clearly endogenous itself, which RS does not seem to deal with.

(2009) and above, there is a sensible theoretical foundation for heterogeneity of impact. A recent study re-examines this issue.

The study by Angeles and Neanidis (2009) (AN) starts with the familiar hypothesis that aid effectiveness likely is hampered by diversion; i.e., aid being diverted away from ideal use, by the ruling elite. In contrast to previous studies, however, AN tries to identify the genesis of the ruling elite, so as to include a measure describing its influence. Specifically, AN identifies the colonial era as the period during which a foreign ruling elite emerged in many development countries. Perhaps especially in the Americas, but also in Africa. The fundamental hypothesis is that immigrants of European decent, after independence, rose to control the political institutions and had limited interest in fostering development for other groups in society. For this to be a credible story the extent of European settlements needs to be sufficiently extensive to enable controlling influence, while at the same time sufficiently limited so as to leave the "Euro descendants" a minority of the population. To fix ideas, AN suggests that places where European descendents make out between 15 and 30 percent of the population strikes this intermediate balance. The paper provides country examples of how the harmful influence from the elite of European decent have led to the waste of aid funds, including Bolivia, El Salvador and Nicaragua. In sum, the hypothesis investigated by AN is whether the impact from aid is smaller in places where Europeans settled to a sufficient extent *during the colonial period*. This is an operational hypothesis in that data for settlements during the 19th century is available.

Their baseline analysis largely follows RS in terms of controls in addition to aid. The key novelty lie in the appearance of an "aid/settlements" interaction term, which is hypothesized to be negative. Another difference to RS lies in the choice of instruments. As an instrument for aid flows, AN proposes an interaction between aid donations at the donor level, and the geographical and cultural distance between the recipient and the individual donors. The cultural distance is measured by variables such as common language, and common religion. Notice that this variable is similar in spirit to that proposed by RS, in that it is thought to capture non-economic determinants of aid allocation to individual countries. In contrast to RS, however, AN does not formally establish a link between (e.g.) geographical distance and actual aid donations.

The key result of AN's analysis is that aid enters linearly with a positive sign, but that aid interacted with European settlers is negative. Both terms are statistically significant. Economically, their results are not too different from those surveyed above. In areas *without* European colonial settlements, we find an elasticity of about 0.24, which is rather close to the RS estimate: An increase in aid (as a fraction of GDP) by 1% should increase long-run

GDP per capita by 0.24%.[29] However, factoring in settlers the *average* impact of aid in their sample is negative.[30]

AN exposes their main findings to a rather demanding battery of robustness checks. In particular they examine whether their interaction term is dominated by the main alternatives proposed in the recent literature (policy, tropical area, etc), and find that the settler interaction survives the simultaneous inclusion of them all. This suggests that their interaction term either complements previous findings of heterogeneity, or, in fact provides a better representation of source of the heterogeneity. Overall their results appear remarkably robust.

Nevertheless, concerns similar to those raised vis-à-vis the RS study applies here. The key external instrument relies on the assumption that the geographical distance (for one thing) between recipient nations and the OECD (i.e., "the donors") *only* influences GDP per capita through foreign aid. This is a questionable assumption. For instance, it is well known that geographical distance between countries (including whether they share borders) is a key determinant of trade patterns (e.g., Frankel and Romer, 1999). It is also a strong determinant for the interaction of citizens via travel, which facilitates technology diffusion (Andersen and Dalgaard, 2009). Since both cross-border flows of goods and people undoubtedly influence GDP per capita directly the key assumption necessary for the results to be valid can be questioned. Being the devil's advocate: The fact that AN is able to establish similar results with alternative strategies may be because these alternatives are equally flawed.

Thus, in spite of creative attempts one may worry that the literature as yet has to convincingly identify the impact of aid. However, the endeavor of identifying the impact of total aid may well be overly ambitious. Aid covers numerous types of interventions which may have separate effects on growth. As a consequence one may feel that a more disaggregated approach would be more promising from the perspective of gaining insights into the impact of aid. In fact a small set of studies have recently started following this track.

---

[29] The authors run extensive robustness checks, and varies the instruments as well. Here we use their own preferred specification, Table 2, column 2, to calculate the economic impact.

[30] This is similar to the conclusion in Dalgaard et al. (2004), where the average impact is negative due to the intervening effect of climate related circumstances.

3.2. Disaggregated studies: The impact of aid modalities

The study that pioneered the examination of disaggregated aid in a cross-country setting is Clemens, Raddlet and Bahvani (2004) (CRB). The key advance in the paper is to examine the influence from a category of aid which CRB argue could hold detectable effects on growth in the short run: "Short term aid".

In this category the authors place budget support (or program aid) as well as project aid given for real sector investments for infrastructure or to directly support production in transportation (including roads), communications, energy, banking, agriculture and industry. This type of aid is thought to be the counterpart of "long-term aid", which includes technical assistance as well as investments in social infrastructure (which encompasses investments in health, schooling etc). If we add humanitarian to short term and long term aid, we are left with the total aid flow.

One challenge for the disaggregation process is that disaggregated disbursements are only available for the 1990s onward. In order to obtain data for disaggregated aid during the 1970s and 1980s the authors make use of commitments. Consequently, the key assumption required for this procedure to be valid is that the composition of commitments is a reasonable proxy for the composition of disbursements. At least for the period during which both commitments and disbursements are available this assumption seems to be a reasonable one. Hence, using the composition of commitments, and actual total disbursements, the authors construct short term aid for the 1970s and 1980s.

In terms of the identification strategy the authors chiefly make use of internal instruments. Consequently, past (disaggregated) aid serve as an instrument for present disaggregated aid. In addition, CRB allow for non-linear effects of aid, along the lines of Hansen and Tarp (2001); short term aid is subject to diminishing returns.

Their estimations reveal an economically (very) sizable and statistically significant impact from short term aid: evaluated at the mean level of short term aid CRB estimates support an elasticity above 2. The main virtue of the study is that it attempts to come to grips with the question of why aid might work. Moreover, at face value it suggests that some forms of aid are in fact highly effective in stimulating growth, which is an interesting finding. At the same time, however, there are reasons why one might choose to be skeptical about the analysis.

First, the elasticity is much larger than what was recovered in previous studies, and several fold larger than what one would expect on *a priori* grounds. This is a cause for concern as the components in "short term aid" largely was assessed in Section 2. If the estimate by CRS is to be supported one will have to re-examine the assumptions made in the calibrations to

assess what would be needed. Possibly aid investments could be subject to considerable external effects, which would serve to elevate the aggregate effect. But it is far from obvious that the required magnitudes could be motivated with strong external validity, since externalities in human and physical capital accumulation have proven hard to detect.[31]

Second, the decomposition of aid raises questions. In particular it is not entirely obvious that program aid is that much different from project aid, which is placed in the "long term aid" category. What if the government chooses to spend the budget support on education, or health? Given the set-up, these forms of spending should then rightly belong to "long term aid" even though they by construction are part of "short term aid". Furthermore, with this in mind, we are faced with difficulties discussed above in the context of identification: it is not clear that government spending in the past, insofar as it is used on human resources, is irrelevant to current growth.

If identification is jeopardized, what do we make of the large estimate for the aid elasticity? The central issue in this regard is whether aid allocation "works" the same way for "short term aid" as it does for total aid. For total aid we know that the poorer the country the more aid per capita it receives. As discussed in Dalgaard and Hansen (2009), this "allocation effect" tends to bias standard (OLS) estimates of aid in a negative direction. But what about budget support, or support for infrastructure investments? It is possible (but it has admittedly not been examined and established in the literature) that the bias could be working in the opposite direction. That is, perhaps more budget support is being donated to more well functioning aid recipients? If this is the case, and if identification fails, the estimates obtained by CRB may be biased in an *upward* direction due to the allocation of budget (or infrastructure) aid.

In spite of these concerns it is clear that CRB opened the door to another angle on the inquiry into aid effectiveness, which already has been taken up by others.

Annen and Kosempel (2009), for instance, explore the impact from *technical assistance* on growth in income per capita. Their central argument is that technical assistance (TA) stimulates human capital accumulation by facilitating knowledge transfer. Accordingly, they disaggregate aid into TA and non-technical assistance. In addition the authors hypothesize that donor fragmentation in a country is harmful to the process of aid-induced human

[31] Moreover, in the short run a large impact from aid on capital accumulation will – perhaps surprisingly – not increase the short run growth impact as it also tends to reduce the size of the gap to the steady state which is closed every period ("the rate of convergence" declines in the terminology of the field). See Dalgaard and Erickson (2009) for details.

capital accumulation, and argue this is particularly the case when it comes to TA. The main result is a TA/GDP per capita elasticity of about 0.2 absent donor fragmentation.

The identification strategy is similar to previous contributions in the literature. They both employ internal and external instruments in an effort to show robustness. In the context of the external instruments the authors follow Rajan and Subramanian in using colonial past, common language, and population ratios (as a proxy for donor influence) to predict TA receipts. Hence, the concerns voiced above vis-à-vis these approaches can also be forwarded here. Moreover, there is conflicting evidence to be found in the literature.

The initial study by CRB, for starters, categorized technical assistance as "long term aid", for which they find only a limited (statistically insignificant) impact on growth. Similarly, a study by Ouattara and Strobl (2008) does not detect any positive impact from technical assistance. Since all three studies employ a similar identification strategy it is not entirely clear what the source of this discrepancy is. It could be due to poor identification, or may be explained by the fragmentation interaction, which neither Quattara and Strobl (2008) nor CRB invoke.

Quattra and Strobel in addition find a positive impact from project aid, which is somewhat inconsistent with CRB. The authors also estimate a *negative* influence from financial program aid, and a zero impact from food aid (which is perhaps unsurprising). Throughout Quattara and Strobl rely in internal instruments with the weaknesses this involve.

Another sort of decomposition is due to Minoiu and Reddy (2009). Rather than focusing on the use of aid (budget support, technical assistance etc) this study innovates in focusing on the motives for aid donations.

It has long been recognized in the allocation literature, that foreign aid is donated for a variety of reasons: both need considerations and geo-political or strategic considerations seems to be at play. Often the latter seems to explain more of the allocation than the former. This observation, as pointed out by the authors, may also matter in the context of aid effectiveness, if the motives for giving aid influence its end use. Concretely Minoiu and Reddy argue that some donors might be delivering more "developmental" aid than others; this could arguably be the case for aid delivered by donors that have recipient *needs* more in focus, rather than those having strategic or perhaps geo-political motivations at heart. Accordingly, they distinguish between donations from "development oriented" donors from aid coming from donors supposedly less preoccupied with development. Recognizing the inherent difficulty in partitioning the donor's into such groups, the authors experiment with

three different but partially overlapping segmentations. The Scandinavian donors are involved in all three groups.

A remarkable finding is that even OLS regressions display a positive correlation between growth and aid from the so-called "like-minded" donors. This finding is surprising in that these donors tend to focus on poverty alleviation (indeed these donors are largely selected this way). Hence, the negative aid allocation bias discussed in Dalgaard and Hansen (2009) should therefore work "full force". And yet, the estimate does not seem to reflect this fact. The association prevails also when the authors invoke instrumental variable estimation, using internal instruments. The estimates suggest a small, but statistically significant, elasticity of about 0.1.

An interesting aspect of the analysis is that authors try to incorporate the fact that aid might have an impact after a considerable time lag; MR are regressing aid donations over the 1970-80 period on growth from 1990-2000. Hence, the complaint that aid works with a time lag (which jeopardizes identification when internal instruments are invoked) is partly dealt with.

Nevertheless, there is reason to be cautious. First, the "lagging of aid procedure" does not seem to yield a positive impact on growth unless growth over the 1990-2000 period is considered. Why should the effect only be visible in the growth record for the 1990s? Moreover, the identification strategy is likely problematic in spite of the time lags. As explained in Section 2 health investments plausibly do not show up until 40 or 60 years down the line. A lag of 10 years is then still not enough to ensure that the exclusion restriction is fulfilled.

3.3. Summary and over-all assessment of the evidence

In spite of a flurry of interest in aggregate aid effectiveness considerable uncertainties and controversy persists. Very little progress has been made in terms of identification since the late 1990s, which is a problem both in terms of understanding the impact of total aid, as well as aid modalities.

Nevertheless, some tentative conclusions do seem to emerge. Table 2 summarizes the results for the studies discussed above. In spite of all the disagreements, few estimates exceed the theoretical priors discussed in Section 2. At best the elasticity of aid with respect to the long run level of income is about 0.3. At the same time, given all the uncertainties discussed above, it is hard to attach much weight to such estimates.

Table 2: Summary of empirical estimates					
STUDY	**AID MEASURE**	**NON-LINEAR EFFECTS**	**INSTRUMENT**	**"BEST CASE" ELASTICITY OF AID WRT GDP PER CAPITA**	**AID STATISTICALLY SIGNIFICANT**
Burnside and Dollar (2000)	Total aid	Yes: Policy	External	0,24%[a]	Yes (with good policies)
Dalgaard et al. (2004)	Total aid	Yes: Area in the tropics	Internal	0,4%[b]	Yes (outside the tropics)
Rajan and Subramanian (2008)	Total aid	No	External	0.3%	No
Angels and Neanidis (2009)	Total aid	Yes: European Settlements in 19th century	External	0.25%[c]	Yes (with few settlers)
Clemens et al. (2004)	Short term aid (budget support, project aid) vs long term aid (e.g., technical assistance)	Yes: Diminishing returns	Internal	2,4%[d]	Yes (short term aid), No (long term aid)
Annen and Kosempel (2009)	Technical assistance	Yes: Donor fragmentation	External and Internal	0,2[e]	Yes (with low fragmentation)
Minoiu and Reddy (2009)	"Developmental aid".	No	Internal	0,1[f]	Yes (developmental), No (other)
Ouattara Strobl (2008)	Project aid, technical assistance, food aid, financial program aid	Yes: various	Internal	…	Yes, positive (project); No (tech assistance, food aid); Yes, negative (financial program)

Notes: The elasticity's are calculated as - b(aid)/b(initial GDP per capita), where the numerator is the marginal impact from aid, and the denominator is the coefficient on initial income, which captures conditional convergence. On theoretical grounds the coefficient on initial GDP per capita is always invoked even if statistically insignificant. *a*) Evaluated at mean policy index, and using estimates from Table 4, column 5. *b*) evaluated outside the topics (Table 3, column 3); *c*) evaluated for zero European settlers (Table 2, column 2). *d*) Evaluated at mean "short term" aid (Table 4, column 5). e) Elasticity of technical assistance per capita wrt to income per capita, assuming zero donor fragmentation and average level of technical assistance (Table 1, column 3). f) Evaluated at mean "group 1" disbursements, estimates from Table 6, column 4.

The literature on aid modalities is still at an early stage. Still, there seems to be some indication that project aid potentially has a larger impact on economic development than other forms of aid. There is also some hints that the identity of the donor (i.e., the aid strategy implicitly invoked) matters. But whether this finding stands up to closer scrutiny (including an improved identification strategy), why it matters, and how it matters is yet unknown.

4. Concluding remarks

In conclusion we offer some perspectives on issues the aid effectiveness literature might sensible try to resolve in the coming years.

There is little doubt that disaggregation should be in focus; aggregate aid flows are simply too diverse to allow for strong conclusions. The literature has already taken some steps in this direction, but only to a somewhat limited extent. It seems important to examine well defined subcomponents of aid, to assess its effectiveness. Does infrastructure aid actually help build infrastructure? That is, does it bring more roads, railways and an improved electricity grid? Does financial program aid support financial development? Do family planning programs affect fertility patterns? None of these questions have been addressed directly in a cross-country setting. They have at best been approached indirectly via the link between broader aid modalities and economic growth. The transmission mechanisms between various kinds of investments and economic outcomes have been left more or less unexplored.

By focusing on more specific areas of aid intervention there might be some hope that identification can be improved. But this requires the careful study of the *allocation* of sub-categories of aid. There is a large allocation literature on aid per capita, but very little work has been done in examining why some countries receive more infrastructure aid than others, why some countries receive support for family planning while others do not, and so forth. Exploring questions such as these might ultimately form the basis for a clearer understanding of how aid influences economic outcomes in the poorest countries on the planet.

Literature

Acemoglu, D. and S. Johnson. , 2007. Disease and development: The effect of life expectancy on economic growth. *Journal of Political Economy* 115(6):925—985.

Andersen, T.B. and C-J. Dalgaard, 2009. Cross-Border Flows of People, Technology Diffusion and Aggregate Productivity. Mimeo (University of Copenhagen)

Annen, K. and S. Kosempel, 2009. Foreign Aid, Donor Fragmentation, and Economic Growth. The B.E. Journal of Macroeconomics: Vol. 9: Iss. 1 (*Contributions*), Article 33.

Angeles, L. and K.C. Neanidis, 2009. Aid Effectiveness: the Role of the Local Elite. *Journal of Development Economics*, 90, 120-34.

Arellano, M. and S.R. Bond, 1991. Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies*, 58, 277-97.

Ashraf Q., A. Lester and D. Weil, 2008. When Does Improving Health Raise GDP? In: Acemoglu, D., K. Rogoff and M. Woodford (eds.) *NBER macroeconomics Annual 2008*. Cambridge M.A.

Bloom, D. and D. Canning and G. Fink, 2009. Disease and Development Revisited. NBER working paper 15137.

Burnside C. and D. Dollar, 2000. Aid, Policies and Growth, *American Economic Review*, 90, 847-68.

Caselli, F. and Feyrer, J., 2007. The Marginal Product of Capital. *Quarterly Journal of Economics,* 122, 535-568.

Caselli, F, G. Esquivel and F. Lefort, 1996. Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics. *Journal of Economic Growth, 1,* 363-389.

Cervellati, M. and U. Sunde, 2009. Life Expectancy and Economic Growth: The Role of the Demographic Transition. Mimeo.

Chanda, A. and C-J. Dalgaard, 2008. Dual Economies and International Total Factor Productivity Differences: Channeling the Impact from Institutions, Trade and Geography. *Economica*, 75, p. 629-61.

Clemens, M., S. Raddlet, and R. Bahvnani, 2004. Counting Chicken's Before the Hatch: The Short Term Effect of Aid on Growth. Center for Global Development Working Paper No. 44.

Chenery, H. B. and A.M. Stout, 1966. Foreign Assistance and Economic Development. *American Economic Review, 56,* 679-733.

Dalgaard, C-J and H. Hansen, 2001. On Aid, Growth and Good Policies. *Journal of Development Studies* 37, 17-41.

Dalgaard, C-J and H. Hansen and F. Tarp, 2004. On the Empirics of Foreign Aid and Growth. *Economic Journal,* 114, 191-216.

Dalgaard C-J., and L. Erickson, 2009. Reasonable Expectations and the First Millennium Development Goal: How Much Can Aid Achieve? *World Development* 37*,* pp. 1170-81

Dalgaard, C-J. and H. Hansen, 2009. Evaluating Aid Effectiveness in the Aggregate: Methodological Issues. *Evaluation Study 2009/01*. Ministry of Foreign Affairs of Denmark.

Deverajan, S., M.J. Miller and E.V. Swanson, 2002. Goals for Development: History, Prospects and Costs. World Bank Policy Research Working Paper No. 2819, World Bank.

Easterly, W., 1999. The Ghost of Financing Gap: Testing the Growth Model Used in the International Financial Institutions. *Journal of Development Economics, 60,* 23-38.

Easterly, W., R. Levine and D. Roodman, 2004. Aid, Policies, and Growth: Comment. *American Economic Review*, 94, 774–80.

Frankel J. and D. Romer, 1999. Does trade cause growth? *American Economic Review*, 89, 379-99

Hansen, H. and F. Tarp, 2001. Aid and Growth Regressions. *Journal of Development Economics*, *64*, 547–70.

Hoeffler, A., 2002. The Augmented Solow Model and the African Growth Debate. *Oxford Bulletin of Economics & Statistics, 64,* 135-58.

Jones, C., 2009. Intermediate Goods and Weak Links: A Theory of Economic Development. Mimeo (Stanford GSB).

Mankiw G., D. Romer and D. Weil, 1992. A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics, 107,* 407–37.

Minoiu, C. and S.G. Reddy, 2009. Development Aid and Economic Growth: A Positive Long-Run Relation. IMF Working Paper 09/118.

Pritchett, L., 1994. Desired fertility and the impact of population policies. *Population and Development Review*.

Ouattara B., and E. Strobl, 2008. Aid, Policy and Growth: Does Aid Modality Matter? *Review of World Economics*, 144, 347-65.

Roodman, D., 2007. The Anarchy of Numbers: Aid, Development, and Cross-Country Empirics. *World Bank Economic Review*, 21, 255-77

Röller, L-H and L. Waverman, 2001. Telecommunications Infrastructure and Economic Development: A Simultaneous Approach. *American Economic Review*, 91, pp. 909-923

Rajan R.G. and A. Subramanian, 2008. Aid and Growth: What Does the Cross-Country Evidence Really Show? *Review of Economics and Statistics*, 90, 643-65.

Selaya, P. and E. Sunesen, 2008. Does Aid Increase Foreign Direct Investment? Mimeo (University of Copenhagen)

Solow, R., 1956. A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics, 70,* 65-94.

Sunesen, E., 2008. Examining the Regional Aspect of Foreign Direct Investment to Developing Countries. Mimeo (University of Copenhagen).

Annex 1. The impact of population growth on labor productivity

Suppose aggregate GDP is produced using the following aggregate production function

$$Y = K^{\alpha}\left(AL\right)^{\beta} X^{\lambda},$$

Where *Y* is GDP, *K* is equipment capital, *L* is the labor force and *X* is land, which is assumed to be fixed over time. Finally, *A* represents technology, which grows at the rate *g*. The exponents parameterizes the share of *K*, *L* and *X* in national accounts, respectively. Constant returns to scale in all inputs (*K*, *L* and *X*) is assumed. Hence $\alpha + \beta + \lambda = 1$.

The production function can be rewritten on the following form:

$$\frac{Y}{L} = \left(A\right)^{\frac{\beta}{1-\alpha}}\left(\frac{K}{Y}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{X}{L}\right)^{\frac{\lambda}{1-\alpha}} = \left(A\right)^{\frac{\beta}{1-\alpha}}\left(\frac{K}{Y}\right)^{\frac{\alpha}{1-\alpha}}\left(\frac{X}{L}\right)^{\frac{\lambda}{\lambda+\beta}}.$$

To quickly get to the formula, we invoke the fact that standard growth models predict that *K/Y* is constant in the long run. We can therefore derive the long run (steady state) growth rate in GDP per worker (labor productivity) by log differentiating the above equation with respect to time. We obtain

$$g_{Y/L} = \frac{b}{1-a}g - \frac{l}{l+b}g_L,$$

where g_L is the growth rate of population. Hence (where Δ signifies absolute changes)

$$\Delta g_{Y/L} = -\frac{\lambda}{\lambda+\beta}\Delta g_L.$$

Since the sum $\lambda + \beta$ is capital's share in national accounts, it follows that If Δg_L is 3 percent, λ is about 0.1, and $\lambda + \beta = 0.4$, the consequence of a lowering of population growth by 3 percent is an increase in $g_{Y/L}$ by about 0.75 percentage points.

Annex 2. Infrastructure and Productivity

We begin with the production function

$$Y = K^{\alpha}\left(AL\right)^{1-\alpha} = K^{\alpha}\left(EG^{\gamma}L\right)^{1-\alpha}$$

Where productivity can be decomposed into a component involving efficiency, *E*, and the influence from infrastructure G^{γ}. Given this production function, Röller and Waverman (2001) (RV) estimate that $\gamma/(1-\alpha)$ is about 0.15.

Now, assume infrastructure is accumulated as follows (time indices suppressed for simplicity)

$$\dot{G} = \sigma Y - \delta_G G\,,$$

where a "dot" denotes a time derivative. The parameter σ is the fraction of output invested in roads, power supply etc, and δ is the rate of capital depreciation. Physical capital is accumulated similarly (Solow, 1956)

$$\dot{K} = sY - \delta_K K\,.$$

Now, abstracting from technological change and population growth, we find that in the long run

$$\lim_{t\to\infty}\left(K/Y\right) = \frac{s}{\delta_K}, \lim_{t\to\infty}\left(G/Y\right) = \frac{\sigma}{\delta_G}$$

$$\lim_{t\to\infty}\left(G/K\right) = \frac{\sigma}{\delta_G}\frac{\delta_K}{s}.$$

Thus, we can write long-run labor productivity

$$\lim_{t\to\infty} Y/L = \left(s/\delta_K\right)^{\frac{\alpha}{1-\alpha}} E\left(\sigma/\delta_G\right)^{\gamma}.$$

Hence a one percent increase in the infrastructure investment rate instigates an increase in long-run productivity of γ percent. According to RV $\gamma/(1-\alpha)$ is about 0.15. Since $(1-\alpha)$ is about 0.67 it follows that γ is about 0.23.

Annex 3. The Input-Output Multiplier and Productivity

This appendix draws on Jones (2009), and represents a toy version of the full framework. Still, it is sufficient for present purposes.

To see how the presence of intermediate goods may lead to a powerful multiplier in dynamic general equilibrium, we begin with the production function

$$Y = A\left(K^{\alpha}L^{1-\alpha}\right)^{\sigma} Z^{1-\sigma}$$

It is important to stress that *Y* here is gross output, not value added (i.e., GDP). GDP in this economy is gross output, net of intermediate goods production. In the above equation, *Z* is intermediate goods, and σ is the share of intermediate goods in gross output.

Total gross output can be used in two basic ways; it can be invested or consumed. Suppose a share *z* is used for investments in intermediate goods (which depreciates fully in production), and a share *s* is used to invest in physical capital *K*. As a result the share (1-*s*-*z*) of output is consumed. In sum

$$Z = zY$$

and

$$\dot{K} = sY - \delta K,$$

Where a "dot" signifies a time derivative, as usual. In the steady state (ignoring technological change, changes in *A*, and population growth, changes in *L*) we have that the capital-output ratio convergence to

$$\frac{K}{Y} = \frac{s}{\delta}.$$

Meanwhile, *Z* is given by (at all points in time)

$$\frac{Z}{Y} = z$$

As a result, output per worker (i.e., GDP per worker is simply (1-*z*) times this expression) in the long run

$$\left(\frac{Y}{L}\right) = \left[Az^{1-\sigma}\right]^{\frac{1}{(1-\alpha)\sigma}}\left(\frac{s}{\delta}\right)^{\frac{\alpha}{1-\alpha}} = \left[Az^{1-\sigma}\left(\frac{s}{\delta}\right)^{\alpha}\right]^{\frac{1}{(1-\alpha)\sigma}}.$$

The key observation is this: if intermediate goods are absent (σ=1), the "multiplier" is the familiar one: $1/(1-\alpha)$. Note that the above formula gives the "rule of thumb" for capital investment in this case, as the exponent for s is $\alpha/(1-\alpha)$, which falls in an interval from ½ to 2/3 depending on the size of α; capital's share in national accounts.

But if $\sigma < 1$, and intermediate goods are present, it is larger. Empirically, since σ is roughly ½ (see Jones, 2009), the multiplier is large. If α is about 0.3 it becomes 3.

As an example, a difference in technology (A) of a factor of two translates into a difference in TFP of a factor of 8. Hence, relatively modest differences in A can generate large country-wide differences in living standards.

The interaction with infrastructure was mentioned in the text. In the context of the production function adopted here the appropriate specification is $A = EG^{g}$, if we are to draw on Röller and Waverman's estimate. This implies that the exponent on G in the expression for long-run Y/L becomes $\gamma/[(1-\alpha)\sigma]$. Since γ is about 0.15, and α is about 1/3, σ is ½, we find that $\gamma/[(1-\alpha)\sigma]$ is about 0.45.

Published by:
Ministry of Foreign Affairs of Denmark
Evaluation Department
Asiatisk Plads 2
1448 Copenhagen K
Denmark
E-mail: eval@um.dk

The publication can be downloaded from:
www.evaluation.dk

ISBN: 978-87-7087-305-5
ISBN: 978-87-7087-306-2 (internet version)

Other Evaluation Studies:
2009/5: Synthesis of Evaluations on Support to Business Development
2009/4: Synthesis of Impact Evaluations of Microcredit
2009/3: Experiences with Conducting Evaluations Jointly with Partner Countries
2009/2: Addressing the Underlying and Basic Causes of Child Undernutrition in Developing Countries: What Works and Why?
2009/1: Evaluating Aid Effectiveness in the Aggregate: Methodological Issues
2008/5: Public Private Partnership Programme
2008/4: Economic Empowerment of Women
2008/3: Assessing Multilateral Organisation Effectiveness
2008/2: Synergies between Bilateral and Multilateral Activities
2008/1: Synthesis of Evaluations of HIV/AIDS Assistance
2007/3: Evaluating the Impact of Rural Roads in Nicaragua
2007/2: Synthesis of Evaluations on Technical Assistance
2007/1: Assistance to Development Research